Rank deficiency in sparse random GF[2] matrices

R.W.R. Darling* Mathew D. Penrose† Andrew R. Wade‡
Sandy L. Zabell§

26th November 2012

Abstract

Let M be a random $m \times n$ matrix with binary entries and i.i.d. rows. The weight (i.e., number of ones) of a row has a specified probability distribution, with the row chosen uniformly at random given its weight. Let $\mathcal{N}(n, m)$ denote the number of left null vectors in $\{0, 1\}^m$ for M (including the zero vector), where addition is mod 2. We take $n, m \to \infty$, with $m/n \to \alpha > 0$, while the weight distribution may vary with n but converges weakly to a limiting distribution on $\{3, 4, 5, \ldots\}$; let W denote a variable with this limiting distribution. Identifying M with a hypergraph on n vertices, we define the *2-core* of M as the terminal state of an iterative algorithm that deletes every row incident to a column of degree 1.

We identify two thresholds α^* and $\underline{\alpha}$, and describe them analytically in terms of the distribution of W. Threshold α^* marks the infimum of values of α at which $n^{-1} \log \mathbb{E}[\mathcal{N}(n, m)]$ converges to a positive limit, while $\underline{\alpha}$ marks the infimum of values of α at which there is a 2-core of non-negligible size compared to n having more rows than non-empty columns.

We have $1/2 \leq \alpha^* \leq \underline{\alpha} \leq 1$, and typically these inequalities are strict; for example when $W = 3$ almost surely, numerics give $\alpha^* = 0.88949\ldots$ and $\underline{\alpha} = 0.91793\ldots$ (previous work on this model has mainly been concerned with such cases where W is non-random). The threshold of values of α for which $\mathcal{N}(n, m) \geq 2$ in probability lies in $[\alpha^*, \underline{\alpha}]$ and is conjectured to equal $\underline{\alpha}$.

The random row weight setting gives rise to interesting new phenomena not present in the non-random case that has been the focus of previous work.

Key words: Random sparse matrix, null vector, hypercycle, random allocation, XORSAT, phase transition, hypergraph core, random equations, large deviations, Ehrenfest model

AMS Subject Classification: 60C05 (Primary) 05C65; 05C80; 15B52; 60B20; 60F10 (Secondary)

*Mathematics Research Group, National Security Agency
†Department of Mathematical Sciences, University of Bath
‡Department of Mathematical Sciences, University of Durham
§Mathematics Department, Northwestern University

Contents

1 Introduction

Suppose that $M := M(n,m)$ is an $m \times n$ matrix with entries in $\{0,1\}$, each of whose rows contains at least one 1, for which we seek a left null vector over GF[2], i.e. a row vector $a \in \{0,1\}^m$ such that $aM \equiv \mathbf{0} \pmod 2$, where here and elsewhere $\mathbf{0}$ is the all-0 vector. More generally, elements of M might belong to the finite field GF[q] of order q. We are interested in the case where M is sparse and random, as specified below.

Let $X_1, X_2, \ldots, X_m$ denote the vectors constituting the rows of M, and let $\sigma(n,m)$ denote the co-rank over GF[2], namely

$$\sigma(n,m) := m - \dim \operatorname{span}\{X_1, X_2, \ldots, X_m\}, \tag{1.1}$$

where here and subsequently 'span' indicates the linear span over GF[2]. Then the number of null vectors of M, including the zero vector, is

$$\mathcal{N}(n,m) = 2^{\sigma(n,m)}, \tag{1.2}$$

which counts the number of distinct solutions in $\{0,1\}^m$, including the zero solution, to

$$a_1 X_1 + \cdots + a_m X_m \equiv \mathbf{0} \pmod 2. \tag{1.3}$$

Note that for a fixed n and a given realization of the sequence of rows $X_1, X_2, \ldots$, the numbers $\mathcal{N}(n,m)$ are nondecreasing as m increases.

Suppose that $n, m \to \infty$, with $m/n \to \alpha > 0$. Our goal is to examine the limiting behaviour of the expected number $\mathbb{E}[\mathcal{N}(n,m)]$ of left null vectors, and the limiting probability $\mathbb{P}[\sigma(n,m) > 0]$ of a mod-2 linear dependency of the rows of $M(n,m)$, as a function of the parameter α, and especially to derive computable thresholds at which phase transitions occur. We also study the rate of exponential decay of the probability that $\mathbf{1} := (1,1,\ldots,1)$ is a null vector.

The probabilistic setting that we consider has the rows $X_1, X_2, \ldots, X_m$ being independent and identically distributed (i.i.d.) with the law of a random vector $X = X(n) \in \{0,1\}^n$. The problem has different flavours depending on the underlying law of X, and several regimes have received considerable attention in the literature, including:

(a) The *dense* regime in which X has order n non-zero components; the standard model studied in this regime has X distributed uniformly over $\{0,1\}^n$.

(b) The classical *sparse* regime in which X has order $\log n$ non-zero components.

(c) The uniformly (very) sparse regime in which X has $O(1)$ non-zero components.

The main focus of the present paper is regime (c) (albeit our '$O(1)$' may be random for each row, and might not even have a mean); in Section 2.7 below we briefly discuss other models that have been studied. In the simplest case, X contains a fixed number $r \le n$ of non-zero components (the cases $r = 1$ and $r = 2$ have distinct behaviour from the case $r \ge 3$); in more generality the number of non-zero components is randomly distributed according to a given *weight distribution*.

Before formally describing our model in detail and presenting our main results (in Section 2), we make some remarks on motivation, and on the literature. Note that $\sigma(n,m) = 0$ if and only if M has row rank m, which occurs if and only if M has column rank m. Thus the absence of non-trivial left null vectors is equivalent to all

column vectors in $\{0,1\}^m$ being expressible as a linear combination of the columns of M (with addition modulo 2), or in other words, to there being a solution $x \in \{0,1\}^n$ to $Mx \equiv y$ for all column vectors $y \in \{0,1\}^m$. In the special case of $r = 2$, motivation for considering this question is discussed at the start of [24, Chapter 3]. The following interpretations help to motivate the general case.

A scheduling problem. Suppose that a tennis club is organizing its annual schedule. There are n playing days, and m potential players. Each player wants to play on a given subset of the days; if there is not a match available on every one of these days, they refuse to pay the annual membership. Each day, in order for nobody to be left out, an even number of players is required. Each possible schedule satisfying these requirements is a left null vector mod 2; the one with the most units achieves the maximal income for the tennis club.

Randomized Lights Out. This is a variant of the game 'Lights Out' [33]. Each of m lamps can be either on or off, and there are n switches, each of which is incident to a specified subset of the lamps, as given by the random matrix M; Lamp i and Switch j are mutually incident if and only if the entry at (i, j) of M is 1. If a switch is toggled, all of the lamps incident to it have their status changed from on to off or off to on. One may ask whether all states (i.e. configurations of on and off lamps) are accessible from the 'all off' state by using some sequence of switches (or equivalently, if the 'all off' state is accessible from all possible starting states), and this is equivalent to the question of whether the column rank of M is m.

The XORSAT problem. This is a variant of the random satisfiability problem [31], where there are n Boolean variables which may be deemed true or false. Each row of M represents a clause built as the logical XOR (exclusive OR) involving those Boolean variables corresponding to columns incident to this row, so the clause is true if an odd number of the variables incident to the row are deemed true. Given a vector $y \in \{0,1\}^m$, finding a solution x to $Mx \equiv y$ corresponds to finding a truth-assignment for the Boolean variables so that each clause i is true if $y_i = 1$ and false if $y_i = 0$. Thus the column rank is m if and only if the problem is satisfiable for all possible choices of y.

A spin-glass model. The relationship between satisfiability problems and spin glasses has already been noted in [31]. In the present instance, consider the following variant of the well-known Sherington–Kirkpatrick mean-field spin-glass model (see e.g. [34]). There is a random collection of hyperedges on n vertices, represented by the m rows of M. Each hyperedge i has a sign g_i, taking value $(-1)^{y_i}$. Each vertex j is assigned a spin σ_j taking values in $\{-1, 1\}$, and (at zero temperature) the probability measure on the state-space is concentrated on states of minimal energy, i.e. with maximal value of $\sum g_i e_i$, where here e_i denotes the product of spins at vertices in hyperedge i. The existence of a configuration with all terms in the sum equal to $+1$ is equivalent to the existence of a solution to $Mx \equiv y$.

The Ehrenfest urn model and the random walk on the hypercube. In the Ehrenfest model of heat exchange, a box contains n particles, some of which are red and the rest blue. At each step, a particle is sampled uniformly from the box and changes its colour. For a sample of the large literature, see e.g. [16, p. 121], [30, §3.5], [32, §3.5], or [21].

In the case where X has a single unit entry, we may view each row of M as selecting which particle is to be changed at that step. Then $\mathbf{1}$ is a null vector for M if and only if the model returns to the initial state after m steps. This may also be interpreted as a random walk on the graph whose vertices are $\{0,1\}^n$ and edges are present between those vertices that differ in a single component; the event that $\mathbf{1}$ is null corresponds to the walker being back in his starting state after m steps.

The general case, allowing other weight distributions, corresponds to a generalization of the Ehrenfest model whereby multiple 'diffusions' are allowed, i.e. at each step several particles may change colour at once; cf [30, Chapter 10]. This can be similarly interpreted in terms of a walk on a version of the hypercube with additional edges.

There is a large body of work on the properties of random matrices over finite fields and the closely related subject of random linear equations over finite fields. Surveys are provided by the book [24, Chapter 3] as well as the articles [28, 29]. The problems may also be formulated in terms of *random hypergraphs*: each row represents a hyperedge, and each column represents a vertex; for details see Section 5 below. They are also related to the *XORSAT* problem in Boolean algebra, as mentioned above (see also Section 2.6.2 below). Generally, such models can be described in the framework of *random allocation* or *occupancy* problems: see the books [21, 24, 26, 30].

The null-vector problem in the fixed row-weight case has received several treatments in the literature, and it is not easy to reconcile all of the existing results, due to differences in presentation and also differences in the underlying probabilistic models. One contribution of the present paper is to clarify some of these issues, including giving a rigorous justification that the results are unchanged under small perturbations of the underlying model. Our main contribution, however, is to treat the case of genuinely *random* row weights, which has not previously been studied. We mention that there has recently been renewed interest in this area in several scientific communities: for example, Alamino and Saad [1] give a statistical physics approach to the null-vector problem; Ibrahimi *et al.* [19] treat the related problem of random XORSAT; Costello and Vu [9] study the rank of random symmetric (so in particular, square) matrices.

Throughout the paper, we extend the function $x \mapsto x^x$, $x > 0$, continuously to $x = 0$, so that $0^0 := 1$. We define the *weight* of a vector $v = (v_1, \ldots, v_n) \in \{0,1\}^n$ to be $w(v) := \sum_{i=1}^n v_i$, i.e., the number of unit entries. For $n \in \mathbb{N} := \{1, 2, \ldots\}$ we write $[n] := \{1, 2, \ldots, n\}$. We write $\xrightarrow{d}$ for convergence in distribution.

2 Results and discussion

2.1 Description of the random matrix model

Given $n \in \mathbb{N}$, suppose that $X = X(n) \in \{0,1\}^n$ is a random row vector, selected according to some probability law on $\{0,1\}^n$. Consider a sequence of i.i.d. random vectors $X_1, X_2, \ldots$ with the same law as X. Let $M := M(n,m)$ be the $m \times n$ matrix whose rows are $X_1, X_2, \ldots, X_m$.

We will consider X with law of the following form. Let W be an $\mathbb{N}$-valued random variable (so $\mathbb{P}[W \geq 1] = 1$) whose law will be the (limiting) *weight distribution* of our random vector X. Let $W_1, W_2, \ldots$ be a sequence of random variables with $W_n \in [n]$ such that $W_n \xrightarrow{d} W$ as $n \to \infty$. Let $w(X)$ have the distribution of W_n, and for

each $k \in [n]$ let the conditional distribution of X, given $w(X) = k$, be uniform over $\{x \in \{0,1\}^n : w(x) = k\}$.

Let $\rho(s) := \mathbb{E}[s^W]$ and $\rho_n(s) := \mathbb{E}[s^{W_n}]$ denote the probability generating functions of W and W_n, respectively. We use $\mathbb{P}_{\rho_n}$ and $\mathbb{E}_{\rho_n}$ for the probability and expectation for the random matrix model with n columns and row weight distribution given by ρ_n. We shall say W_n are *uniformly bounded* if there is a finite constant r_1 such that $\mathbb{P}[W_n \leq r_1] = 1$ for all n (and hence $\mathbb{P}[W \leq r_1] = 1$ as well).

2.2 Threshold results in the general setting

Given the probability generating function ρ, define the threshold

$$\alpha^*_\rho := \inf\{\alpha \geq 0 : F_\rho(\alpha) > 0\}, \tag{2.1}$$

where we set

$$F_\rho(\alpha) := \log \sup_{\gamma \in [0,1/2]} \left(\frac{(1 + \rho(1-2\gamma))^\alpha}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right), \quad \alpha \geq 0. \tag{2.2}$$

We state some fundamental properties of $F_\rho(\alpha)$ and α^*_ρ in the next result, which we prove in Section 4.2.

Proposition 2.1. *We have $F_\rho(\alpha) = 0$ for $0 \leq \alpha \leq \alpha^*_\rho$ but $F_\rho(\alpha) > 0$ for $\alpha > \alpha^*_\rho$, and F_ρ is continuous and nondecreasing as a function of α. Moreover:*

(i) *$\alpha^*_\rho \in [0,1]$; and $\alpha^*_\rho < 1$ if $\mathbb{E}[W] < \infty$.*

(ii) *$\alpha^*_\rho = 0$ if and only if $\mathbb{P}[W = 1] > 0$.*

(iii) *If $\mathbb{P}[W = 2] = 1$, then $\alpha^*_\rho = 1/2$.*

(iv) *Suppose that $\tilde{W}$ is another $\mathbb{N}$-valued random variable, with $\tilde{\rho}(s) = \mathbb{E}[s^{\tilde{W}}]$, such that $\tilde{\rho}(s) \leq \rho(s)$ for all $s \in [0,1]$ (which is the case, for example, if $\tilde{W}$ stochastically dominates W). Then $\alpha^*_{\tilde{\rho}} \geq \alpha^*_\rho$.*

In particular, if $\mathbb{P}[W \geq 2] = 1$ and $\mathbb{E}[W] < \infty$, then $\alpha^*_\rho \in [1/2, 1)$.

Here is our first main result, describing the threshold behaviour of the expected number of null vectors. We shall prove this in Section 4.5.

Theorem 2.2. *Suppose that $m_n/n \to \alpha \in (0, \infty)$ as $n \to \infty$. Then*

$$\lim_{n \to \infty} n^{-1} \log \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n)] = F_\rho(\alpha); \tag{2.3}$$

in particular, the limit in (2.3) is strictly positive for $\alpha > \alpha^*_\rho$ and zero for $\alpha \leq \alpha^*_\rho$. Moreover, if in addition there exist $r_0 \geq 3$ and $r_1 < \infty$ such that $\mathbb{P}[r_0 \leq W_n \leq r_1] = 1$ for all n, and $\alpha \in (0, \alpha^*_\rho)$, then as $n \to \infty$,

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n)] = 1 + O(n^{2-r_0}). \tag{2.4}$$

Remark. For $\alpha < \alpha^*_\rho$, the expectation in (2.4) is dominated by low-weight null vectors (short hypercycles), in particular by null vectors with only 2 non-zeros. It may be possible, by extending the argument of Lemma 4.3 below, to expand this expectation as a power series in n^{-1}, but we do not pursue this here. For $\alpha > \alpha^*_\rho$ the exponential rate in (2.3) is dominated by null vectors using a specific positive proportion $\beta_0 = \beta_0(\alpha)$ of the (roughly αn) available rows (and also possibly those using proportion $1 - \beta_0$ of the rows, due to parity effects); in fact, $\beta_0 = \frac{\rho(1-2\gamma_0)}{1+\rho(1-2\gamma_0)} \in (0, 1/2)$, where $\gamma_0 = \gamma_0(\alpha) \in (0, 1/2)$ is the value of γ for which the supremum in (2.2) is attained. See Section 4.5 for details.

Theorem 2.2 deals with the *expected* number of null vectors. Also of interest is, for a fixed n, the (random) number of rows m at which the first non-zero null vector appears. Define

$$T_n := \min \{m \in \mathbb{N} : X_m \in \operatorname{span} \{X_1, X_2, \ldots, X_{m-1}\}\}, \tag{2.5}$$

the first m for which $\operatorname{rank}(M(n, m)) < m$. Standard linear algebra implies that $T_n \leq n+1$.

We define another threshold, $\underline{\alpha}_\rho$, through an analytic description that needs more notation (we shall give a probabilistic interpretation later on). For $x \in (0, 1)$ set

$$\psi(x) := x + \left(1 + \frac{\rho(x)}{\rho'(x)} - x\right) \log(1 - x); \tag{2.6}$$

$$h(x) := -\frac{\log(1 - x)}{\rho'(x)}. \tag{2.7}$$

Provided $\mathbb{P}[W \geq 1] = 1$ we can and do extend ψ continuously to $\psi(0) := 0$, since $\rho(s)/\rho'(s) = O(s)$ as $s \downarrow 0$. Note that $h(x) \to \infty$ as $x \downarrow 0$ provided $\mathbb{P}[W \geq 3] = 1$, and that if $\mathbb{E}[W] < \infty$ then as $x \uparrow 1$ we have $h(x) \to \infty$ and $\psi(x) \to -\infty$. Set

$$\alpha^\sharp_\rho := \inf_{x \in (0,1)} h(x), \tag{2.8}$$

and note that $\alpha^\sharp_\rho \rho'(x) \leq -\log(1 - x)$ for all $x \in (0, 1)$, so integrating from 0 to 1 we get $\alpha^\sharp_\rho \leq 1$, provided $\mathbb{P}[W \geq 1] = 1$. For $\alpha \geq 0$, define

$$g^*(\alpha) := \sup\{x \in (0, 1) : h(x) \leq \alpha\}, \tag{2.9}$$

with the convention $\sup \emptyset = 0$ in operation in (2.9). Observe that if h has unbounded range (e.g. if $\mathbb{P}[W \geq 3] = 1$) then $h \circ g^*$ is the identity map on $[\alpha^\sharp_\rho, \infty)$. See Figure 1 for an example. Define

$$\underline{\alpha}_\rho := \inf\{\alpha > \alpha^\sharp_\rho : \psi(g^*(\alpha)) < 0\}. \tag{2.10}$$

In (2.10), the set defining $\underline{\alpha}_\rho$ is non-empty provided $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, since as $\alpha \to \infty$ we have $g^*(\alpha) \to 1$ and $\psi(g^*(\alpha)) \to -\infty$.

The relevance of $\underline{\alpha}_\rho$ for the null vector problem is shown by the next result.

Theorem 2.3. *Suppose W_n are uniformly bounded and $\mathbb{P}[W_n \geq 3] = 1$ for all n. Then $\alpha^*_\rho \leq \underline{\alpha}_\rho \leq 1$, and for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}_{\rho_n} \left[(\alpha^*_\rho - \varepsilon)n \leq T_n \leq (\underline{\alpha}_\rho + \varepsilon)n\right] = 1. \tag{2.11}$$

Theorem 2.3 is proved in Section 5.3. We exclude the case where $\mathbb{P}[W \in \{1, 2\}] > 0$ from the statement of Theorem 2.3; different phenomena occur in that case (see Proposition 2.9 below). This case is also discussed in [10], where the functions ψ and h also play a role.

Typically, $\underline{\alpha}_\rho$ defined by (2.10) will satisfy $\psi(g^*(\underline{\alpha}_\rho)) = 0$. In many cases, there is a single solution in $(0,1)$, denoted x^*_ρ, to $\psi(x) = 0$, and $\underline{\alpha}_\rho = h(x^*_\rho)$. However, the situation is complicated by the fact that $\alpha \mapsto g^*(\alpha)$ typically has at least one discontinuity. We defer a more detailed discussion of the properties of the functions ψ, h, and g^*, and the corresponding thresholds, to Section 5.3 below. Figure 1 provides an example.

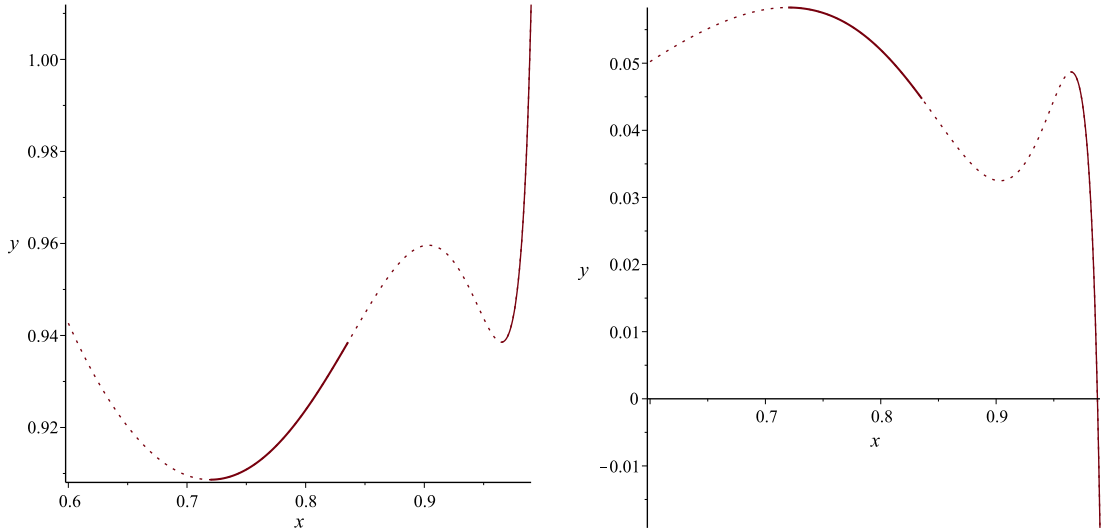

Figure 1: Example with $\rho(s) = 0.9s^3 + 0.1s^{24}$. The left plot shows parts of the curves $y = h(x)$ (all the line) and $x = g^*(y)$ (solid line). The right plot shows parts of the curves $y = \psi(x)$ (all the line) and the locus of $(g^*(\alpha), \psi(g^*(\alpha)))$ (solid line). The left plot shows that $g^*(\alpha)$ has two discontinuities, one at $\alpha = \alpha^\sharp_\rho \approx 0.908654$ and one at $\alpha \approx 0.938536$, with the first corresponding to a jump from $g^* = 0$ to $g^* \approx 0.719682$ and the second to a jump from $g^* \approx 0.835696$ to $g^* \approx 0.964919$. The right plot shows the single positive solution of $\psi(x) = 0$ at $x = x^*_\rho \approx 0.987817$, so $\underline{\alpha}_\rho = h(x^*_\rho) \approx 0.991613$. It is not a coincidence that the curves h and ψ seem to mirror each other: see Lemma 5.9 below.

Theorem 2.3 leaves open the sharp asymptotics of T_n/n: we believe that the upper bound in Theorem 2.3 (i.e., $\underline{\alpha}_\rho$) is sharp:

Conjecture 2.4. *If W_n are uniformly bounded and $\mathbb{P}[W_n \geq 3] = 1$ for all n, then T_n/n converges in probability to $\underline{\alpha}_\rho$ as $n \to \infty$.*

An equivalent statement to the fixed-weight case $W = r \geq 3$ of this conjecture seems to have been established recently in the random-XORSAT literature: see the comments in Section 2.6.2.

The probabilistic interpretation of the thresholds $\alpha^\sharp_\rho$ and $\underline{\alpha}_\rho$ is in terms of the *2-core* of $M(n,m)$; this is the terminal state of an iterative algorithm that deletes every row incident to a column of degree 1 (see Section 5 below for details). Let $E(n,m;\varepsilon)$ denote the event that $M(n,m)$ possesses a 2-core (i) whose number of rows is bounded below by εn, and (ii) which contains more rows than columns of non-zero weight (all of which have weight 2 or more). In particular, for $\varepsilon > 0$, $E(n,m;\varepsilon)$ implies that $M(n,m)$ has a non-empty 2-core. If W_n are uniformly bounded, and $m_n/n \to \alpha > 0$, then in Theorem 5.6 below we will show that, under certain additional conditions on ρ, there exists $\varepsilon > 0$ (allowed to depend on α) such that $\lim_{n\to\infty} \mathbb{P}_{\rho n}[E(n, m_n; \varepsilon)] = 1$ for α in some interval of the form $(\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$ with $\delta > 0$.

A more delicate description of the behaviour of the 2-core in terms of the function ψ defined at (2.6) will be given in Theorem 5.6 below: the limiting aspect ratio of the 2-core being less than or greater than 1 depends on the sign of $\psi(g^*(\alpha))$. In the example shown in Figure 1, and also in the fixed weight setting, $\psi(g^*(\alpha))$ changes sign only once, but in the the general random weight setting it may change sign *multiple* times; see the example in Figure 3 below. Thus the random weight setting gives rise to subtle new phenomena not present in the fixed weight case that has been the focus of previous work.

Remark. Our techniques may be used to obtain information about the *weight profile* of the left null space. Bayes' Theorem applied to equation (4.2) below shows that the weight of a randomly chosen left null vector has distribution given by

$$\mathbb{P}_{\rho_n}[w(v) = k \mid vM \equiv \mathbf{0}] = \frac{\binom{m}{k}\mathbb{P}_{\rho_n}[A(n,k)]}{\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]};$$

the quantities on the right-hand side here are studied in detail in Section 3.

2.3 Even occupancy in random allocations

Let $A(n,m)$ denote the event that the row vector $\mathbf{1} = (1,\dots,1)$ is null for M, i.e.,

$$A(n,m) := \{X_1 + \cdots + X_m \equiv \mathbf{0} \pmod 2\}. \tag{2.12}$$

One interpretation is in terms of the *random allocation model*. Suppose we have n urns, and for each row of M we allocate a collection of balls to a chosen set of urns (determined by the unit entries of that row of M). Event $A(n,m)$ is the event that all the urns end up with an even number of balls. Random allocations have been extensively studied; see e.g. [16, p. 101], and the monographs [21, 24, 26, 30].

The following theorem, which we prove in Section 3, describes the exponential rate of decay for $\mathbb{P}_{\rho_n}[A(n,m_n)]$ where m_n/n has a finite positive limit. The theorem excludes the case in which *both* W and m_n only take odd values; if m is odd and W_n is odd a.s., then $\mathbb{P}_{\rho_n}[A(n,m)] = 0$ since the total number of units in the matrix is odd.

Theorem 2.5. *Suppose that $m_n/n \to \alpha \in (0,\infty)$ as $n \to \infty$, and that either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Then*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}_{\rho_n}[A(n,m_n)] = -R_\rho(\alpha), \tag{2.13}$$

where $R_\rho(\alpha) > 0$ is continuous and nondecreasing in $\alpha > 0$ and is defined by

$$R_\rho(\alpha) := -\log \sup_{\gamma\in[0,1/2]} \left(\frac{(\rho(1-2\gamma))^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}} \right). \tag{2.14}$$

The relevance of Theorem 2.5 to Theorem 2.2 is clear; see the formula (4.3) below. We present an interesting special case of Theorem 2.5 that can be understood in isolation. Let $\pi_n(m)$ denote the probability that all the components Y_j of a multinomial $(m; n^{-1},\dots,n^{-1})$ random vector $(Y_1,\dots,Y_n)$ are even. Here Y_j can be interpreted as the occupancy of urn j after m balls are independently and uniformly distributed into n distinct urns: see e.g. [32, p. 23], [30, p. 11], or [21, p. 90].

Then $\pi_n(m) = 2^{-n}\sum_{j=0}^n \binom{n}{j}(1-(2j/n))^m$; this formula is known in the Ehrenfest urn literature (see [30, pp. 128–129]) and can also be obtained from (3.1) below. If m is odd, $\pi_n(m)$ must be zero.

Proposition 2.6. *Let $\pi_n(m_n)$ denote the probability that all the n components of a multinomial $(m_n; n^{-1}, \dots, n^{-1})$ random vector are even. Suppose that m_n is even for each n and $m_n/n \to \alpha = \lambda \tanh \lambda \in (0, \infty)$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} n^{-1} \log \pi_n(m_n) = \log \cosh \lambda - (\lambda \tanh \lambda)(1 - \log \tanh \lambda). \tag{2.15}$$

This result follows from Theorem 2.5, as we shall show in Section 3.5. Proposition 2.6 can also be derived from a result of Kolchin [23, Theorem 2, p. 141].

2.4 The fixed row-weight case

We now describe the special case of the results in Section 2.2 when $\mathbb{P}[W = r] = 1$, for some fixed r. The existing literature is largely concerned with this case (see the discussion in Section 2.6 below).

Let $r \in \mathbb{N}$. Define the thresholds α_r^*, $\alpha_r^\sharp$, and $\underline{\alpha}_r$ to be the values of α_ρ^*, $\alpha_\rho^\sharp$, and $\underline{\alpha}_\rho$, respectively, in the case where $\mathbb{P}[W = r] = 1$ (i.e. where $\rho(s) = s^r$). By (2.1) we have

$$\alpha_r^* := \inf\{\alpha \geq 0 : F_r(\alpha) > 0\},$$

where

$$F_r(\alpha) := \log \sup_{\gamma\in[0,1/2]} \left(\frac{(1 + (1-2\gamma)^r)^\alpha}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right). \tag{2.16}$$

By Proposition 2.1, $F_r(\alpha) = 0$ for $\alpha \leq \alpha_r^*$ but $F_r(\alpha) > 0$ for $\alpha > \alpha_r^*$. If $r \geq 3$, $\psi(x) = 0$ has a single solution in $(0, 1)$ (see Proposition 5.8 below) denoted x_r^* and satisfying

$$x_r^* = -\left(1 - \left(\frac{r-1}{r}\right) x_r^*\right) \log(1 - x_r^*), \tag{2.17}$$

and

$$\underline{\alpha}_r = h(x_r^*) = -\frac{\log(1 - x_r^*)}{r(x_r^*)^{r-1}}. \tag{2.18}$$

For example, $\alpha_3^\sharp \approx 0.818469$, $g^*(\alpha_3^\sharp) \approx 0.715332$, and $x_3^* \approx 0.883414$ so $\underline{\alpha}_3 \approx 0.917935$ (see also Figure 4 below). The next result is a specialization of Theorems 2.2 and 2.3 together with Theorem 5.6 and Proposition 5.7.

Theorem 2.7. *Let $r \in \mathbb{N}$. Suppose that $W_n \to r$ in probability. Suppose that $m_n/n \to \alpha \in (0, \infty)$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n)] = F_r(\alpha); \tag{2.19}$$

in particular, the limit in (2.19) is strictly positive if and only if $\alpha > \alpha_r^*$. If also $\mathbb{P}[W_n \geq 3] = 1$ (so $r \geq 3$) and W_n are uniformly bounded, then $\alpha_r^* \leq \underline{\alpha}_r \leq 1$ and for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_{\rho_n} \left[(\alpha_r^* - \varepsilon)n \leq T_n \leq (\underline{\alpha}_r + \varepsilon)n \right] = 1,$$

and moreover, there exists $\varepsilon_\alpha > 0$ such that for all $\varepsilon \in (0, \varepsilon_\alpha)$,

$$\lim_{n\to\infty} \mathbb{P}_{\rho_n}[E(n, m_n; \varepsilon)] = \begin{cases} 0 & \text{if } \alpha < \underline{\alpha}_r \\ 1 & \text{if } \alpha > \underline{\alpha}_r. \end{cases} \tag{2.20}$$

r	1	2	3	4	5	6	7	8
$\alpha_r^\sharp$	0	0.5	0.818469	0.772280	0.701780	0.637081	0.581775	0.534997
α_r^*	0	0.5	0.889493	0.967147	0.989162	0.996228	0.998650	0.999510
$\underline{\alpha}_r$	—	—	0.917935	0.976770	0.992438	0.997380	0.999064	0.999660

Table 1: Threshold parameters for r-uniform random hypergraphs. Note that $\underline{\alpha}_r$ is not defined when $r = 1$ or 2.

The sharp monotone transition displayed by (2.20) in the fixed weight case was indicated by Cooper [8]. In the general, random weight setting, the picture can be more complicated, and there exist examples where the transition is *non-monotone*: see the example in Figure 3 below.

Remark. In the case $r = 2$, compare Theorem 2.7 to Proposition 2.9(ii) below: the first cycle in the random graph appears at $m_n = Zn$, where Z has an asymptotic distribution on $(0, 1/2)$. It is a classical result that for $\alpha \in (0, 1/2)$, if $m_n/n \to \alpha$, the number of cycles $\mathcal{N}(n, m_n)$ in an Erdős–Rényi graph has a Poisson limit with finite expectation, but the limiting expectation is infinite for $\alpha \geq 1/2$ (see e.g. [24, §2.3]); we could not find in the literature an explicit reference to the fact that the expectation blows up *exponentially* with n for $\alpha > 1/2$, at the rate given by Theorem 2.7 (a classical result of Erdős and Rényi states that at $\alpha = 1/2$, the expected number of cycles grows as $\frac{1}{4} \log n$: see Theorem 5a of [15, p. 41]).

2.5 Threshold numerics and asymptotics

In this section we discuss numerical and asymptotic evaluation of the thresholds in our results for the fixed row-weight case described in Section 2.4. Table 1 shows values of $\alpha_r^\sharp$, α_r^*, and $\underline{\alpha}_r$, for $r \leq 8$. Previous computations of these thresholds are reviewed in Section 2.6. As suggested by the numerical results, it can be shown that, for r large enough, $\alpha_r^\sharp < \alpha_r^* < \underline{\alpha}_r < 1$; this is a consequence of the following result.

Proposition 2.8. *As $r \to \infty$,*

$$\alpha_r^\sharp \to 0; \quad 1 - \alpha_r^* \sim \frac{\mathrm{e}^{-r}}{\log 2}; \quad 1 - \underline{\alpha}_r \sim \mathrm{e}^{-r}. \tag{2.21}$$

The asymptotic result for α_r^* in (2.21) is due to Calkin [5]; we prove the other two parts in Section 5.3 below.

One can obtain arbitrarily sharp upper and lower bounds for the solution $x_r^* \in (0, 1)$ of $\psi(x) = 0$ in the case $\rho(s) = s^r$, $r \geq 3$, as follows. In this case, by (2.17) we have that $x_r^* = i_r(x_r^*)$, where we set

$$i_r(x) := 1 - \exp\left\{ -\frac{x}{1 - (\frac{r-1}{r})x} \right\}.$$

For $\theta \in [0, 1)$, $x \mapsto \frac{x}{1-\theta x}$ is strictly increasing for $x \in [0, 1]$. Thus if $x_r^* > a_n$, it follows that $x_r^* > a_{n+1} := i_r(a_n)$. Also, $i_r'(0) = 1$, $i_r''(0) = \frac{2-r}{r} > 0$, and $i_r(1) = 1 - \mathrm{e}^{-r} < 1$, so $i_r(x) > x$ for $x \in (0, x_r^*)$ but $i_r(x) < x$ for $x \in (x_r^*, 1]$. Hence starting with $a_0 = \frac{r-2}{r-1} < x_r^*$ (an inequality proved in Proposition 5.8 below), we can iterate to obtain an increasing sequence of lower bounds a_n for x_r^*. Conversely, starting instead with $b_0 = 1 > x_r^*$

and iterating $b_{n+1} := i_r(b_n)$ gives a decreasing sequence of upper bounds b_n for x_r^*. For example, after one step we get

$$1 - \exp\left\{-\frac{r(r-2)}{2(r-1)}\right\} < x_r^* < 1 - \mathrm{e}^{-r}, \quad (r \geq 3).$$

Proceeding up to b_2 for the upper bound and a_4 for the lower bound is sufficient to obtain the $r \to \infty$ asymptotic expression

$$x_r^* = 1 - \mathrm{e}^{-r} - r^2\mathrm{e}^{-2r} + O(r^4\mathrm{e}^{-3r}), \tag{2.22}$$

which will be the main ingredient in the proof of the $\underline{\alpha}_r$ result in (2.21).

In fact, this iterative procedure converges, so $a_n \uparrow x_r^*$ and $b_n \downarrow x_r^*$. To prove convergence it is sufficient to show that $i_r'(x) < 1$ at $x = x_r^*$. A calculation shows that $i_r'(x)$ evaluated at $x = x_r^*$ comes to $x^{-2}(1-x)(\log(1-x))^2$, so for the required inequality it suffices to show that

$$-x^{-1}\log(1-x) < (1-x)^{-1/2}, \text{ for } 0 < x < 1.$$

The coefficient of x^k in the power series expansion of the left-hand side of the last inequality is $1/(k+1)$, and for the right-hand side it is $4^{-k}\binom{2k}{k}$, and both series are convergent on the given interval. An induction shows that $1/(k+1) \leq 4^{-k}\binom{2k}{k}$ for all integers $k \geq 0$. So term-by-term comparison of the two power series gives the inequality.

To end this section we discuss the numerical evaluations of α_r^* in Table 1. In this discussion we use some claimed properties of the functions involved that we do not verify rigorously, since here we are only concerned with numerical estimation. Let

$$F_{r,\alpha}(\gamma) := \log\left(\frac{(1+(1-2\gamma)^r)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right), \tag{2.23}$$

so that $F_r(\gamma) = \sup_{\gamma\in[0,1/2]} F_{r,\alpha}(\gamma)$. Differentiating, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}F_{r,\alpha}(\gamma) = -\frac{2\alpha r(1-2\gamma)^{r-1}}{1+(1-2\gamma)^r} + \log\left(\frac{1-\gamma}{\gamma}\right). \tag{2.24}$$

Thus γ is a stationary value for $F_{r,\alpha}$ if γ solves

$$\alpha = \frac{1+(1-2\gamma)^r}{2r(1-2\gamma)^{r-1}}\log\left(\frac{1-\gamma}{\gamma}\right) =: \alpha_r(\gamma). \tag{2.25}$$

Numerical curve sketching shows that (2.25) generically has at most 2 solutions in $(0, 1/2)$; of such solutions, the smallest will be the local maximum, since $F_{r,\alpha}'(\gamma) \to \infty$ as $\gamma \downarrow 0$, by (2.24). If (2.25) has no solutions in $(0, 1/2)$, then the supremum in (2.16) is either $F_{r,\alpha}(0) = (\alpha - 1)\log 2$ or $F_{r,\alpha}(1/2) = 0$. Thus setting $\gamma_0 := \gamma_0(\alpha, r) = 0$ if (2.25) has no solutions in $(0, 1/2)$ and $\gamma_0 := \gamma_0(\alpha, r)$ to be the smallest positive solution to (2.25) otherwise, we have that $F_r(\alpha) = F_{r,\alpha}(\gamma_0)$ whenever $\alpha \in (0, 1)$.

For $\alpha \in (0, 1)$, $F_r(\alpha) > 0$ if and only if $\gamma_0(\alpha, r) > 0$. Moreover, Proposition 2.1 shows that $\alpha_r^* < 1$, so that for $\alpha < 1$ such that $\gamma_0(\alpha, r) > 0$, $F_r(\alpha) = \alpha_r(\gamma_0)\log(1+(1-2\gamma_0)^r) - \log(2\gamma_0^{\gamma_0}(1-\gamma_0)^{1-\gamma_0})$. Thus to find α_r^*, we solve for $\gamma \in [0, 1/2]$ the equation

$$\alpha_r(\gamma) - \phi_r(\gamma) = 0, \tag{2.26}$$

where

$$\phi_r(\gamma) = \frac{\log\left(2\gamma^\gamma(1-\gamma)^{1-\gamma}\right)}{\log\left(1+(1-2\gamma)^r\right)}.$$

Numerical curve plotting shows that $\gamma \mapsto \alpha_r(\gamma) - \phi_r(\gamma)$ is decreasing on $[0, 1/2]$, so (2.26) can be solved using efficient numerical methods; let γ_r denote the solution to (2.26). Then we compute α_r^* via $\alpha_r^* = \alpha_r(\gamma_r)$.

2.6 Discussion and related results

2.6.1 Previous results on threshold values

In the simplest case, $W_n = W_n^{\mathrm{hyp}} := r \wedge n$ a.s., for a fixed $r \in \mathbb{N}$; then $W = r$ a.s. This fixed row weight 'hypergeometric' model is studied by Cooper [6]. A variation is the model in which r units are assigned to the row *independently* and uniformly at random, with multiplicities reduced mod 2. The latter 'binomial' model corresponds to $W_n = W_n^{\mathrm{bin}}$ distributed as the number of odd components in a multinomial $(r; n^{-1}, \ldots, n^{-1})$ random vector; then $W_n \xrightarrow{d} r$ (see Lemma 3.3 below). The $r \geq 3$ binomial model is studied by Kolchin [23]. Note that in this model rows of all zeroes may appear, in which case they are ignored (in other words, empty hyperedges are discounted): this is a small effect since $\mathbb{P}[W_n^{\mathrm{bin}} = 0] = O(n^{-r/2})$, so a vanishing proportion of rows needs to be discarded.

Phase transitions in the null vector problem for random matrices over finite fields with fixed row weight $r \geq 3$ have been studied since the early 1990s. In the case of the binomial model, the threshold α_r^*, $r \geq 3$, for $\mathbb{E}[\mathcal{N}(n, m_n)]$, $m_n/n \to \alpha$, was described by Balakin *et al.* [3] and Kolchin [23] (having been announced in [22]); in these results α_r^* is characterized by the fact that the expected number of non-trivial null vectors tends to 0 (∞) when $\alpha < \alpha_r^*$ ($\alpha > \alpha_r^*$), but the proofs show that the growth is in fact exponential for $\alpha > \alpha_r^*$. Calkin [5] and Cooper [6] also study α_r^*, $r \geq 3$, and in particular Calkin [5] studies α_r^* as $r \to \infty$; both [6] and [5] work in the case $W_n = r \wedge n$. Note that Cooper's [6] expression of the matrix problem is transposed compared to ours. The first part of our Theorem 2.7 represents a slight generalization of the results just mentioned above because it allows for any class of sequences W_n provided $W_n \to r$ in probability. The case of finite fields of order $q \geq 3$ has also been studied: see for instance [4, 6, 25].

In these previous investigations, the analytic description of the threshold α_r^* varies. Calkin [5, §4] gives the same description of α_r^* as our equation (2.16). Systems of nonlinear equations for computing α_r^* have been proposed in [3, 6, 23]; these descriptions can be shown to be consistent with ours. Specifically, with $F_{r,\alpha}(\gamma)$ as given at (2.23), one may characterize α_r^* by the two equations $F_{r,\alpha}(\gamma) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}\gamma} F_{r,\alpha}(\gamma) = 0$. On the substitution $\lambda = \frac{1}{2}\log(\frac{1-\gamma}{\gamma})$, the first of these equations becomes, after some calculations along the lines of those in the proof of Proposition 2.6 below,

$$(1 + (\tanh\lambda)^r)^\alpha \mathrm{e}^{-\lambda \tanh\lambda} \cosh\lambda = 1. \tag{2.27}$$

The second equation, involving the vanishing of the derivative given at (2.24) gives, after the same substitution for λ,

$$r\alpha = (1 + (\tanh\lambda)^{-r})\lambda \tanh\lambda. \tag{2.28}$$

The system of equations (2.27) and (2.28) is the same as that given by Cooper [6, p. 269], and, after some manipulation, is seen to coincide also with that given by Balakin

et al. [3, p. 564] and Kolchin [23, p. 139]. Despite this agreement, there are some small discrepancies in the numerical evaluations for α_r^* in [3, 5, 6, 23], which can presumably be put down to numerical inaccuracies.

A similar tabulation to our tabulation of $\underline{\alpha}_r$ is given by Cooper [8, pp. 370–371], who also gives an equivalent analytic description of $\underline{\alpha}_r$ to our (2.17); see also Dietzfelbinger *et al.* [13] which we discuss further in the next subsection. We note also that $\alpha_r^\sharp$ has received considerable attention in its own right: see e.g. [19] for its role in random XORSAT.

2.6.2 Between the two thresholds

The following problem arises in the XORSAT literature. Let $r \in \mathbb{N}$ with $r \geq 3$. Let M be our $m \times n$ matrix, with $m/n \to \alpha > 0$, and suppose $W_n = r$ a.s. for all $n \geq r$. Let N denote the number of column vectors $x \in \{0,1\}^n$ such that $Mx \equiv \omega$, where $\omega \in \{0,1\}^m$ is chosen uniformly at random (independent of M). Thus N is a random variable.

Dubois and Mandler [14] show for $r = 3$, and Dietzfelbinger *et al.* [13] extend to general $r \in \mathbb{N}$ with $r \geq 3$ (also providing a more detailed proof) the following result (see [14, Theorem 3.1] and [13, Theorem 1], and also [19]): there is a constant $\tilde{\alpha}_r > 0$ such that provided $\alpha < \tilde{\alpha}_r$, $\mathbb{P}[N > 0] \to 1$ as $n \to \infty$.

The proof of this in [14] is based on a second moment calculation. The analytical definition of $\tilde{\alpha}_r$ in [13, 14] is not obviously the same as our definition of $\underline{\alpha}_r$, but the definition in terms of cores (see [13, Proposition 3 and equation (4)]) seems to match our definition of $\underline{\alpha}_r$, and the numerical values in [13] are consistent with our $\underline{\alpha}_r$.

If we accept that $\tilde{\alpha}_r = \underline{\alpha}_r$, this result implies that if $\alpha < \underline{\alpha}_r$ there is, for n large enough, no non-zero left null vector for M, as follows. Suppose that a non-zero y satisfies $y \cdot M = \mathbf{0}$. Then $N > 0$ implies $y \cdot \omega = 0$. So $\mathbb{P}[y \cdot \omega = 0] \geq \mathbb{P}[N > 0] \to 1$, which contradicts the easy observation that $\mathbb{P}[y \cdot \omega = 0] = 1/2$ for non-zero y.

We may then deduce that in the case with $W_n = n \wedge r$, our Theorem 2.3 may be strengthened to $n^{-1}T_n \to \underline{\alpha}_r$ in probability. This implies that for α in the interval $(\alpha_r^*, \underline{\alpha}_r)$, a form of substantialism occurs; existence of any left null vector is unlikely, but if there is one, there are lots of them.

2.7 Results for other random matrix models

2.7.1 The case of fixed weight vectors with $r = 1$ or $r = 2$

The classical cases of the constant weight model $W_n = n \wedge r$ in which $r \in \{1, 2\}$ exhibit different behaviour from the case $r \geq 3$. Recall the definition of T_n from (2.5).

Proposition 2.9. *(i) For $r = 1$, for any $z > 0$, $\lim_{n\to\infty} \mathbb{P}[T_n > zn^{1/2}] = \exp\{-z^2/2\}$.*
(ii) For $r = 2$, for any $z \in (0, 1)$,

$$\lim_{n\to\infty} \mathbb{P}[T_n > zn/2] = (1-z)^{1/2} \exp\left\{\frac{z}{2} + \frac{z^2}{4}\right\}. \tag{2.29}$$

Proof. In the case $r = 1$, $X_1, X_2, \ldots, X_m$ correspond to m repetitions of the experiment of placing a ball uniformly at random in one of n urns. Then T_n is the first trial at which a ball is placed in an urn which is already occupied, and its law is given by solution to the birthday problem [16, p. 33]:

$$\log \mathbb{P}[T_n > m] = \sum_{j=1}^{m-1} \log\left(1 - \frac{j}{n}\right) = -\frac{m(m-1)}{2n} + O(m^3/n^2),$$

provided $m = o(n)$. Take $m = zn^{1/2}$ to obtain the result.

In the case $r = 2$, T_n is the same as the number of edges which have to be added in the Erdős–Rényi random graph process in order for the first cycle to appear. The formula (2.29) has been derived by Janson [20, Theorem 8.1]. □

2.7.2 Uniform non-zero random vectors over a finite field

For the sake of comparison with sparse matrix phenomena, the case where X is selected uniformly at random from $\{0,1\}^n \setminus \{\mathbf{0}\}$ is worthy of mention.

In fact we give a result for the more general finite field GF[q] for arbitrary prime power q. Let $E_q^n := \{0, \dots, q-1\}^n$. Suppose (for this section only) that $X_1, X_2, \dots$ are i.i.d. random vectors that are uniformly distributed over $E_q^n \setminus \{\mathbf{0}\}$, and let $M(n,m)$ be the $m \times n$ matrix over GF[q] with rows $X_1, \dots, X_m$. We write $\mathbb{P}_*$ for probability associated with this model. Define T_n by (2.5) (with addition mod q). The following elementary result proves that $n + 1 - T_n$ is likely to be a small integer, uniformly in n. Write $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$.

Proposition 2.10. *Suppose that $X_1, X_2, \dots$ are independent and uniformly distributed on $E_q^n \setminus \{\mathbf{0}\}$. As $n \to \infty$, $n + 1 - T_n$ converges to a distribution on $\mathbb{Z}_+$: for any $r \in \mathbb{Z}_+$,*

$$\mathbb{P}_*[T_n > n + 1 - r] \to \prod_{j=r}^{\infty} (1 - q^{-j}) =: p_r \in [0,1], \tag{2.30}$$

where $p_0 = 0$ and $p_r \sim \exp\{-q^{1-r}/(q-1)\}$ as $r \to \infty$. Moreover, the following lower bounds apply: for $n \in \mathbb{N}$ and $1 \le r \le n$,

$$\mathbb{P}_*[T_n > n + 1 - r] \ge \begin{cases} \exp\{-q^{1-r}\} & \text{if } q \ge 3 \\ \exp\{-\frac{4}{3} 2^{1-r}\} & \text{if } q = 2 \end{cases}.$$

Remark. Limit results of the form of (2.30) are classical, and apparently date back at least to an 1895 paper of Landsberg (see [29, p. 69]). The $q = 2$ case of (2.30) corresponds to the $T = n - s$, $m + s = 0$ case of [24, Theorem 3.2.1, p. 126], but with a slightly different probabilistic model: there the X_i are uniform on the whole of E_2^n, not just $E_2^n \setminus \{\mathbf{0}\}$; comparing the results shows that this difference in the probability measures used is negligible in the limit.

Proof of Proposition 2.10. Let A_k denote the event that $\{X_1, X_2, \dots, X_k\}$ is linearly independent. If A_k occurs, then the span of $X_1, X_2, \dots, X_k$ is a subspace with $q^k - 1$ non-zero elements. Since X_{k+1} is (statistically) independent of $X_1, X_2, \dots, X_k$ and uniform on $E_q^n \setminus \{\mathbf{0}\}$, which has $q^n - 1$ non-zero elements,

$$\mathbb{P}_*\left[A_{k+1} \mid A_k\right] = \frac{q^n - q^k}{q^n - 1},$$

so that, since $\mathbb{P}_*[A_1] = 1$, for $k \in \mathbb{N}$,

$$\mathbb{P}_*[T_n > k] = \mathbb{P}_*[A_k] = \prod_{j=1}^{k-1} \mathbb{P}_*[A_{j+1} \mid A_j] = \prod_{j=1}^{k-1} \frac{q^n - q^j}{q^n - 1} = (1 - q^{-n})^{1-k} \prod_{j=1}^{k-1} (1 - q^{j-n});$$

with the usual convention that an empty product is 1. Taking $k = n + 1 - r$ we obtain

$$\mathbb{P}_*[T_n > n + 1 - r] = (1 - q^{-n})^{r-n} \prod_{j=r}^{n-1} (1 - q^{-j}) \to \prod_{j=r}^{\infty} (1 - q^{-j}), \tag{2.31}$$

on letting $n \to \infty$, establishing (2.30) with p_r as stated there. The asymptotic form given for p_r follows from writing $p_r = \exp \sum_{j=r}^{\infty} \log(1 - q^{-j})$ and applying Taylor's theorem.

Moreover, it follows from the fact that $(1 - q^{-n})^{r-n} \geq 1$ provided $r \leq n$ that the convergence in (2.31) is in fact monotone for $n \geq r$, i.e.,

$$\log \mathbb{P}_*[T_n > n + 1 - r] \geq \sum_{j=r}^{n-1} \log \left(1 - q^{-j}\right) \geq \sum_{j=r}^{\infty} \log \left(1 - q^{-j}\right).$$

With $f(x) = \log(1-x) + x + x^2$, we have $f(0) = 0$ and $f'(x) = \frac{x(1-2x)}{1-x} \geq 0$ for $|x| \leq 1/2$, so $\log(1 - x) \geq -x - x^2$ for all x with $|x| \leq 1/2$. Thus for $q \geq 2$ and $r \geq 1$,

$$\log \mathbb{P}_*[T_n > n + 1 - r] \geq -\sum_{j=r}^{\infty} q^{-j} - \sum_{j=r}^{\infty} q^{-2j} = -\frac{q^{1-r}}{q-1} \left(1 + \frac{q^{1-r}}{q+1}\right).$$

The stated lower bounds follow. □

2.7.3 Random vectors of weight $O(\log n)$

There is a separate class of results, initiated by the classical work of Kovalenko [27] and Balakin [2], in which the weight is random with a law which changes as n increases; in the classical case it is $\mathrm{Bin}(n, (a + \log n)/n)$ for $a \in \mathbb{R}$. Much more on this model is given in [7, 24, 28, 29], for example.

3 Multinomial parities and random allocations

3.1 Overview and terminology

In this section we work towards proving Theorem 2.5 and Proposition 2.6. Our null vector problem can be naturally formulated in terms of classical occupancy problems of random allocations of balls into urns.

We shall use the following terminology. Suppose W is a random variable taking values in $\mathbb{Z}_+$, and $k \in \mathbb{N}$, and $p, p_1, p_2, \ldots, p_k$ are numbers in $[0, 1]$ such that $\sum_{i=1}^{k} p_i = 1$. (In most of the rest of the paper we assume $W \geq 1$, but for this section we can allow W to take value 0.) Let us say the random variable X has the $\mathrm{Bin}(W, p)$ distribution if for each $n \in \mathbb{Z}_+$ the conditional distribution of X, given that $W = n$, is binomial with parameters (n, p). Let us say that a random vector $(Z_1, \ldots, Z_k)$ has the multinomial $(W; p_1, \ldots, p_k)$ distribution if for each $n \in \mathbb{Z}_+$ the conditional distribution of $(Z_1, \ldots, Z_k)$, given that $W = n$, is multinomial with parameters $(n; p_1, \ldots, p_k)$.

Recall from Section 2.1 that we assume W_n (having the distribution of row weights for our matrix with n columns) is chosen to converge in distribution to a limiting random variable W. An important special case is the so-called *binomial* model. In the binomial scheme take $W_n = W_n^{\mathrm{bin}}$ to be distributed as the number of odd components in a multinomial $(W; n^{-1}, \ldots, n^{-1})$ random vector. By Lemma 3.3 below, $W_n^{\mathrm{bin}} \xrightarrow{d} W$ as $n \to \infty$, so this is indeed a special case.

We write $\mathbb{P}^{\text{bin}}_{\rho_n}$ for probability associated with the binomial allocation scheme. For the general model we write $\mathbb{P}_{\rho_n}$ as before.

3.2 Exact formulae for the allocation problem

In this subsection n is fixed. Let X_{ij} denote the jth component of X_i. Define the column sums Y_j and partial row sums $S_{i,J}$ of the matrix (X_{ij}) as follows (in this case the addition does not need to be mod 2):

$$Y_j := \sum_{i=1}^{m} X_{ij},\ j \in [n]; \quad \text{and} \quad S_{i,J} := \sum_{j \in J} X_{ij},\ J \subseteq [n].$$

Recall from (2.12) that $A(n,m)$ denotes the event that $\mathbf{1}$ is a null (row) vector for M.

Lemma 3.1. *In the binomial allocation scheme, we have the exact formula*

$$\mathbb{P}^{\text{bin}}_{\rho_n}[A(n,m)] = 2^{-n} \sum_{j=0}^{n} \binom{n}{j} \left(\rho(1-(2j/n))\right)^m . \tag{3.1}$$

In the general allocation scheme,

$$\mathbb{P}_{\rho_n}[A(n,m)] = 2^{-n} \sum_{J \subseteq [n]} \left(\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}]\right)^m \tag{3.2}$$

$$= 2^{-n} \sum_{j=0}^{n} \binom{n}{j} (2p_j^{(n)} - 1)^m, \tag{3.3}$$

where $p_j^{(n)} := \sum_{r=0}^{n} p_{j,r} \mathbb{P}[W_n = r]$ *and* $p_{j,r}$ *is given by*

$$p_{j,r} = \frac{1}{\binom{n}{j}} \left(\binom{n-r}{j} + \binom{r}{2}\binom{n-r}{j-2} + \binom{r}{4}\binom{n-r}{j-4} + \cdots \right). \tag{3.4}$$

Proof. Event $A(n,m)$ occurs if and only if all the Y_j are even, so

$$\mathbb{P}[A(n,m)] = \mathbb{E} \prod_{j=1}^{n} \left(\frac{1+(-1)^{Y_j}}{2} \right) = 2^{-n} \sum_{J \subseteq [n]} \mathbb{E}\left[(-1)^{\sum_{j \in J} Y_j}\right],$$

where the latter sum is over subsets J of $[n]$, including the empty set. Since $\sum_{j \in J} Y_j = \sum_{i=1}^{m} S_{i,J}$ and $S_{1,J}, S_{2,J}, \ldots$ are i.i.d., (3.2) follows.

Consider the binomial allocation scheme. In the binomial model,

$$S_{1,J} = \sum_{j \in J} X_{1j} \equiv \sum_{j \in J} Z_j \ (\text{mod } 2),$$

where $(Z_1, \ldots, Z_n)$ has a multinomial $(W; n^{-1}, \ldots, n^{-1})$ distribution so that $\sum_{j \in J} Z_j$ has a $\text{Bin}(W, |J|/n)$ distribution. Recalling that if $\xi \sim \text{Bin}(n,p)$ then $\mathbb{E}[s^\xi] = (sp + (1-p))^n$, we then obtain (3.1) from (3.2).

In the general scheme, conditional on $\sum_{j=1}^{n} X_{1j} = r$, the distribution of $S_{1,J}$ is hypergeometric with parameters $(n; |J|, r)$. We do not use the generating function (see

e.g. [32, p. 17]) explicitly, but apply the hypergeometric probability mass function directly to (3.2).

Let $H \subseteq [n]$ denote the set of values of j for which $X_{1j} = 1$. For $r \in [n]$, we write $\mathbb{E}_r$ for expectation in the case where $\mathbb{P}[W_n = r] = 1$. Instead of fixing $J \subseteq [n]$ and choosing H as a uniform random r-subset, we obtain an exact formula for $\mathbb{E}_r[(-1)^{S_{1,J}}]$ by fixing $j = |J|$ and a r-subset H, and selecting J uniformly from the j-subsets of $[n]$. The probability $p_{j,r}$ that $S_{1,J} := |H \cap J|$ is even is given by summing probabilities for $|H \cap J| \in \{0, 2, 4, \ldots\}$, giving the expression in (3.4). It follows that $\mathbb{E}_r[(-1)^{S_{1,J}}] = 2p_{j,r} - 1$, and hence

$$\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}] = \sum_{r=1}^{n} (2p_{j,r} - 1)\mathbb{P}[W_n = r] = 2p_j^{(n)} - 1.$$

Substitution of this into (3.2) gives (3.3). □

3.3 Asymptotics in the binomial model

The remaining parts of Section 3 are concerned with asymptotic analysis of the quantities in Lemma 3.1. First we state a result that will enable us to work primarily with even m, which has technical advantages.

Lemma 3.2. *Suppose that $W_n \xrightarrow{d} W$ and $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Then for any $m > 3$,*

$$\log \mathbb{P}_{\rho_n}[A(n, m-3)] + O(\log n) \leq \log \mathbb{P}_{\rho_n}[A(n, m)] \leq \log \mathbb{P}_{\rho_n}[A(n, m+3)] + O(\log n).$$

Proof. The fact that $\mathbb{P}[W \in 2\mathbb{Z}] > 0$ and $W_n \xrightarrow{d} W$ implies that there exist $\varepsilon > 0$ and $r \in 2\mathbb{Z}$ such that $\mathbb{P}[W_n = r] > \varepsilon$ for all n large enough. For any $m > 3$, suppose that $A(n, m-3)$ occurs. Then $A(n, m)$ will occur if the 3 additional rows themselves constitute a hypercycle. With probability at least ε^3, these new rows each have r units, and given this, there is a probability at least n^{-2r}, say, that these units form a hypercycle. In other words, $\log \mathbb{P}_{\rho_n}[A(n, m)] \geq \log \mathbb{P}_{\rho_n}[A(n, m-3)] + O(\log n)$. Applying this inequality twice, once with $m + 3$ in place of m, gives the result. □

Recall that in general we assume $W_n \xrightarrow{d} W$. Next we give an elementary lemma that confirms the binomial model's place in this framework.

Lemma 3.3. *For W a $\mathbb{Z}_+$-valued random variable, let W_n^{bin} be the number of odd components in a multinomial $(W; n^{-1}, \ldots, n^{-1})$ random vector. Then $W_n^{\mathrm{bin}} \xrightarrow{d} W$ as $n \to \infty$.*

Proof. Let W and W_n^{bin} be coupled in the natural way. Then for each $k \in \mathbb{N}$, by the union bound

$$\mathbb{P}[W_n^{\mathrm{bin}} \neq W \mid W = k] \leq \binom{k}{2} \frac{1}{n},$$

which tends to zero as $n \to \infty$, and the result follows easily from this. □

We will prove Theorem 2.5 (in Section 3.5) by first showing that (2.13) holds in the binomial setting, using (3.1) and the Stirling approximation for the binomial coefficients as discussed in Section 6. Then we will extend this to the general setting using an approximation argument described in Section 3.4. We start by proving a slightly more general statement than (2.13) in the binomial case, which we will also need later in the proof of Theorem 2.2.

Lemma 3.4. *Recall the definition of R_ρ from (2.14). Then $R_\rho(\alpha)$ is continuous and nondecreasing as a function of α. Suppose that either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Suppose that there exist α_1, α_2 with $0 < \alpha_1 < \alpha_2 < \infty$ such that, for all n sufficiently large, $\alpha_1 < m_n/n < \alpha_2$. Then,*

$$\begin{aligned} \limsup_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] &\le -R_\rho(\alpha_1); \\ \liminf_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] &\ge -R_\rho(\alpha_2). \end{aligned} \tag{3.5}$$

In particular, if $m_n/n \to \alpha > 0$, then

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] = -R_\rho(\alpha). \tag{3.6}$$

Proof. Suppose that $m_n/n \in (\alpha_1, \alpha_2)$. First assume that the m_n are even. By (3.1) and Lemma 6.3(iii),

$$\begin{aligned} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m)] &\le (n+1)2^{-n} \max_{0\le j\le n} \binom{n}{j} |\rho(1 - (2j/n))|^m \\ &\le (n+1)2^{-n} \sup_{\gamma\in[0,1/2]} \binom{n}{\gamma n} (\rho(1-2\gamma))^m, \end{aligned}$$

where we set $\binom{n}{x} = 0$ if x is a not an integer in $\{0, 1, \dots, n\}$. Using the upper bound on binomial coefficients from the first inequality in (6.3), we get

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] \le (n+1) \sup_{\gamma\in[0,1/2]} \left(2\gamma^\gamma (1-\gamma)^{1-\gamma}\right)^{-n} (\rho(1-2\gamma))^{m_n}.$$

By monotonicity, we then obtain

$$n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] \le n^{-1} \log(n+1) + \log \sup_{\gamma\in[0,1/2]} g_{m_n/n}(\gamma), \tag{3.7}$$

where we have set

$$g_\alpha(\gamma) := \frac{(\rho(1-2\gamma))^\alpha}{2\gamma^\gamma (1-\gamma)^{1-\gamma}};$$

so with R_ρ as defined at (2.14), $R_\rho(\alpha) = -\log \sup_{\gamma\in[0,1/2]} g_\alpha(\gamma)$. Note that $g_\alpha(\gamma)$ is continuous in γ for $\alpha \ge 0$ and $\gamma \in [0, 1/2]$, and $g_\alpha(\gamma)$ is nonincreasing in α; this monotonicity implies, by Dini's theorem, that if $\alpha' \to \alpha$ monotonically then $g_{\alpha'}$ converges uniformly to g_α on the compact interval $[0, 1/2]$. It follows that $\alpha \mapsto \sup_{\gamma\in[0,1/2]} g_\alpha(\gamma)$ is continuous as a function of $\alpha > 0$, and is also nonincreasing in α. In particular, this shows that $R_\rho(\alpha)$ is continuous and nondecreasing in α, as claimed in the lemma. Moreover, we obtain from (3.7) and the fact that $m_n/n > \alpha_1$ that

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] \le \log \sup_{\gamma\in[0,1/2]} g_{\alpha_1}(\gamma),$$

which gives the first inequality in (3.5).

For the second inequality, we have from (3.1) and (6.4) that for any integer $i_n \le n/2$,

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, m_n)] \ge 2^{-n} \binom{n}{i_n} (\rho(1 - (2i_n/n)))^{m_n} \ge \mathrm{e}^{-1/6} \left(\frac{n}{2\pi i_n (n - i_n)}\right)^{1/2} (g_{m_n/n}(i_n/n))^n,$$

using the fact that m_n is even. Then, since $m_n/n < \alpha_2$,

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \geq \mathrm{e}^{-1/6}\left(\frac{n}{2\pi i_n(n-i_n)}\right)^{1/2}(g_{\alpha_2}(i_n/n))^n.$$

Now use continuity of g_α to choose a sequence of integers $i_n \leq n/2$, $n \in \mathbb{N}$, such that $g_{\alpha_2}(i_n/n) \to \sup_{\gamma\in[0,1/2]} g_{\alpha_2}(\gamma)$, with $i_n \to \infty$ and $n - i_n \to \infty$ as $n \to \infty$. The lower bound in (3.5) follows, for m_n even.

The results in (3.5) extend to the case of odd m_n with $\mathbb{P}[W \in 2\mathbb{Z}] > 0$ by Lemma 3.2, which is applicable here by Lemma 3.3. Finally, (3.6) follows from (3.5) on taking $\alpha_1 = \alpha - \varepsilon$ and $\alpha_2 = \alpha + \varepsilon$, for arbitrary $\varepsilon > 0$, and using the continuity of R_ρ. □

3.4 Approximation by the binomial model

The exact formula (3.1) is simpler to work with than the more complicated exact formula (3.3), but intuition suggests that the asymptotics of any of the models in the class with $W_n \xrightarrow{d} W$ should be similar. In this section we quantify this intuition.

Lemma 3.5. *Suppose that $W_n \xrightarrow{d} W$ and either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Suppose that there exist α_1, α_2 with $0 < \alpha_1 < \alpha_2 < \infty$ and $n_0 \in \mathbb{N}$ such that $\alpha_1 < m_n/n < \alpha_2$ for all $n \geq n_0$. Then, uniformly over sequences m_n satisfying the given conditions,*

$$\lim_{n\to\infty} n^{-1}\left|\log \mathbb{P}_{\rho_n}[A(n,m_n)] - \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\right| = 0. \tag{3.8}$$

In particular, if $m_n/n \to \alpha > 0$,

$$\lim_{n\to\infty} n^{-1}\log \mathbb{P}_{\rho_n}[A(n,m_n)] = \lim_{n\to\infty} n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)].$$

Proof. Denote the weight of row 1 by W_n in the general allocation scheme, and by W_n^{bin} in the binomial scheme. Then $W_n \xrightarrow{d} W$ and $W_n^{\mathrm{bin}} \xrightarrow{d} W$, and we can work in a probability space where $\mathbb{P}[W_n \neq W_n^{\mathrm{bin}}] \to 0$. As in Section 3.2, set $S_{1,J} = \sum_{j\in J} X_{1j}$; on $\{W_n = W_n^{\mathrm{bin}}\}$, the (conditional) law of $S_{1,J}$ is the same as in the binomial model. Thus

$$\sup_{J\subseteq[n]}\left|\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}] - \mathbb{E}^{\mathrm{bin}}_{\rho_n}[(-1)^{S_{1,J}}]\right| \leq 2\mathbb{P}[W_n \neq W_n^{\mathrm{bin}}] \to 0. \tag{3.9}$$

First suppose that the m_n are even. By (3.2) with (3.1) and (3.9), there exists a triangular array of numbers $(\delta_{j,n}, j \in [n]\cup\{0\}, n \in \mathbb{N})$ satisfying $\max_{0\leq j\leq n}|\delta_{j,n}| \to 0$ as $n \to \infty$, and

$$\mathbb{P}_{\rho_n}[A(n,m_n)] = 2^{-n}\sum_{j=0}^{n}\binom{n}{j}\left(\rho(1-(2j/n)) + \delta_{j,n}\right)^{m_n}. \tag{3.10}$$

Let $\varepsilon > 0$ and choose $K > 1$ large enough so that $\log(1-K^{-1}) > -\varepsilon/\alpha_2$ and $\log(1+K^{-1}) < \varepsilon/\alpha_2$. Then choose $\delta > 0$ such that $(K+1)\delta < \exp\{-1/(\alpha_1\varepsilon)\}$. Finally assume n is large enough so that $\sup_{j\in[n]\cup\{0\}}|\delta_{j,n}| \leq \delta$ and $\alpha_1 < (m_n/n) < \alpha_2$.

We split the sum in (3.10) into two parts, depending on the size of $\rho(1-(2j/n))$. First suppose that $|\rho(1-(2j/n))| \leq K\delta$. In this case

$$|(\rho(1-(2j/n)) + \delta_{j,n})^{m_n}| \leq ((K+1)\delta)^{m_n} \leq \exp\{-m_n/(\alpha_1\varepsilon)\} \leq \exp\{-n/\varepsilon\}, \tag{3.11}$$

and similarly,

$$(\rho(1-(2j/n)))^{m_n} \le \exp\{-n/\varepsilon\}. \tag{3.12}$$

It follows from (3.10) and (3.11) that

$$\begin{aligned}\mathbb{P}_{\rho_n}[A(n,m_n)] = 2^{-n} &\sum_{j:|\rho(1-(2j/n))|>K\delta} \binom{n}{j} (\rho(1-(2j/n))+\delta_{j,n})^{m_n} \\ &+ O\left(\exp\{-n/\varepsilon\}\right). \end{aligned} \tag{3.13}$$

Now suppose that $|\rho(1-(2j/n))| > K\delta$. In this case

$$(\rho(1-(2j/n))+\delta_{j,n})^{m_n} = (\rho(1-(2j/n)))^{m_n}\left(1+\theta_{j,n}K^{-1}\right)^{m_n},$$

where $|\theta_{j,n}| \le 1$. By the choice of K, $\mathrm{e}^{-\varepsilon/\alpha_2} < 1+\theta_{j,n}K^{-1} < \mathrm{e}^{\varepsilon/\alpha_2}$, and hence

$$\exp\{-\varepsilon n\} < (1+\theta_{j,n}K^{-1})^{m_n} < \exp\{\varepsilon n\}.$$

Therefore

$$(\rho(1-(2j/n))+\delta_{j,n})^{m_n} = (\rho(1-(2j/n)))^{m_n}\exp\{\varepsilon_{j,n}n\},$$

where $|\varepsilon_{j,n}| < \varepsilon$. Hence for the sum on the right-hand side of (3.13), there exists ε_n with $|\varepsilon_n| < \varepsilon$ such that

$$\begin{aligned}&2^{-n} \sum_{j:|\rho(1-(2j/n))|>K\delta} \binom{n}{j} (\rho(1-(2j/n))+\delta_j)^{m_n} \\ &= 2^{-n}\exp\{\varepsilon_n n\} \sum_{j:|\rho(1-(2j/n))|>K\delta} \binom{n}{j} (\rho(1-(2j/n)))^{m_n},\end{aligned}$$

using the assumption that m_n is even so all the terms in the sum are nonnegative. Then by (3.12) and a similar argument to (3.13), the last displayed quantity is equal to $\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\exp\{\varepsilon_n n\} + O\left(\exp\{(\varepsilon-\varepsilon^{-1})n\}\right)$. Combining this with (3.13) we obtain

$$\mathbb{P}_{\rho_n}[A(n,m_n)] = \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\exp\{\varepsilon_n n\} + O\left(\exp\{(\varepsilon-\varepsilon^{-1})n\}\right), \tag{3.14}$$

uniformly in n (and m_n), the implicit constants depending on α_1 and α_2.

It follows from (3.14) that

$$\log \mathbb{P}_{\rho_n}[A(n,m_n)] = \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] + \varepsilon_n n + \log\left(1 + \frac{\Delta_n}{\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\exp\{\varepsilon_n n\}}\right),$$

where $\Delta_n = O\left(\exp\{(\varepsilon-\varepsilon^{-1})n\}\right)$ is the final term in (3.14). By Lemma 3.4, we have that $\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \ge \exp\{-nR_\rho(\alpha_2) - \varepsilon n\}$, for all n large enough. So we may take $\varepsilon > 0$ small enough so that

$$\log\left(1 + \frac{\Delta_n}{\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\exp\{\varepsilon_n n\}}\right) = O(\exp\{-n\}),$$

say. Hence

$$n^{-1}\log \mathbb{P}_{\rho_n}[A(n,m_n)] = n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] + \varepsilon_n + o(1).$$

Since $|\varepsilon_n| \le \varepsilon$ and $\varepsilon > 0$ was arbitrary, (3.8) follows in the case of even m_n. In the other case, Lemma 3.2 yields the same conclusion. The final statement in the lemma then follows from Lemma 3.4, which says that $\lim_{n\to\infty} n^{-1}\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]$ exists in $(0,\infty)$ when $m_n/n \to \alpha > 0$. □

3.5 Proofs of Theorem 2.5 and Proposition 2.6

Now we can complete the proofs of Theorem 2.5 and Proposition 2.6.

Proof of Theorem 2.5. The theorem is now a consequence of Lemmas 3.4 and 3.5. □

Proof of Proposition 2.6. Set

$$f_\alpha(\gamma) := \log\left(\frac{(1-2\gamma)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right). \tag{3.15}$$

In the case $\rho(s) = s$, it follows from Theorem 2.5 that $n^{-1}\log\pi_n(m_n) \to \sup_{\gamma\in[0,1/2]} f_\alpha(\gamma)$. Proposition 2.6 will follow once we prove that, setting $\alpha = \lambda\tanh\lambda$,

$$\sup_{\gamma\in[0,1/2]} f_\alpha(\gamma) = -(\lambda\tanh\lambda)(1-\log(\tanh\lambda)) + \log(\cosh\lambda). \tag{3.16}$$

Note that $f_\alpha(\gamma) \to -\infty$ as $\gamma \uparrow 1/2$. Differentiating (3.15) gives, for $\gamma \in (0,1/2)$,

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} f_\alpha(\gamma) = -\frac{2\alpha}{1-2\gamma} + \log\left(\frac{1-\gamma}{\gamma}\right),$$

which is zero at $\gamma_1 := \gamma_1(\alpha) \in (0,1/2)$ defined implicitly in terms of α by

$$\alpha = \frac{1}{2}(1-2\gamma_1)\log\left(\frac{1-\gamma_1}{\gamma_1}\right). \tag{3.17}$$

One can verify that for $\alpha > 0$ (3.17) defines a unique stationary value $\gamma_1 \in (0,1/2)$ which is a local maximum, since for $\gamma_1 \in (0,1/2)$ the right-hand side of (3.17) is positive, continuous, and strictly decreasing as a function of γ_1, vanishing at $\gamma_1 = 1/2$; this local maximum is indeed the maximum of $f_\alpha(\gamma)$ for $\gamma \in [0,1/2]$ since $f'_\alpha(\gamma) \to \infty$ as $\gamma \downarrow 0$ (and also $f''_\alpha(\gamma_1) < 0$).

Setting $\lambda = \frac{1}{2}\log\left(\frac{1-\gamma_1}{\gamma_1}\right)$ we see that $\lambda\tanh\lambda = \alpha$ as given by (3.17), since we get $\tanh\lambda = 1-2\gamma_1$. To verify (3.16) we need to express $f_\alpha(\gamma_1)$ in terms of λ to get the expression on the right-hand side of (3.16). We have

$$\begin{aligned} f_\alpha(\gamma_1) &= \alpha\log(1-2\gamma_1) - \log 2 - \gamma_1\log\gamma_1 - (1-\gamma_1)\log(1-\gamma_1) \\ &= (\lambda\tanh\lambda)\log\tanh\lambda + \log\cosh\lambda + ((1/2)-\gamma_1)\log\gamma_1 + (\gamma_1-(1/2))\log(1-\gamma_1), \end{aligned}$$

where we have used the fact that $\log\tanh\lambda = \log(1-2\gamma_1)$ and $\log\cosh\lambda = -\frac{1}{2}\log(1-\tanh^2\lambda) = -\log 2 - \frac{1}{2}\log\gamma_1 - \frac{1}{2}\log(1-\gamma_1)$. Collecting the terms involving γ_1 in the last displayed equation, we see that they simplify to $-\lambda\tanh\lambda = -\alpha$ as given by (3.17), so we verify (3.16). □

3.6 Alternative proof of Proposition 2.6 via Poissonization

We give an alternative proof of Proposition 2.6 based on a Poissonization device (as used by Kolchin in his proof of Theorem 2 in [23]) and large deviations arguments of a slightly different flavour from those in the proof above. The proof in this section is direct, avoiding the general Theorem 2.5, but does use instead some relatively deep local limit theory. The following result can be found for example in [23].

Lemma 3.6. *Suppose that $Z_1, Z_2, \ldots$ are independent Poisson random variables with mean $\mu > 0$. Let $Z_1^E, Z_2^E, \ldots$ be i.i.d., where the law of Z_1^E is the same as the conditional law of Z_1, given that Z_1 is even: $\mathbb{P}[Z_1^E = k] = \mathbb{P}[Z_1 = k \mid Z_1 \in 2\mathbb{Z}]$. Then*

$$\pi_n(m) = \frac{\mathbb{P}\left[Z_1^E + \cdots + Z_n^E = m\right]}{\mathbb{P}\left[Z_1 + \cdots + Z_n = m\right]} \left(\frac{1 + \mathrm{e}^{-2\mu}}{2}\right)^n. \tag{3.18}$$

Proof. By the well-known relationship between the Poisson and multinomial distributions (see e.g. [23, p. 140] or [24, p. 15]), and Bayes' Theorem,

$$\begin{aligned}\pi_n(m) &= \mathbb{P}\left[Z_1 \in 2\mathbb{Z}, \ldots, Z_n \in 2\mathbb{Z} \mid Z_1 + \cdots + Z_n = m\right] \\ &= \frac{\mathbb{P}[Z_1 + \cdots + Z_n = m \mid Z_1 \in 2\mathbb{Z}, \ldots, Z_n \in 2\mathbb{Z}]}{\mathbb{P}[Z_1 + \cdots + Z_n = m]} \cdot \mathbb{P}[Z_1 \in 2\mathbb{Z}]^n,\end{aligned}$$

which with the expression for $\mathbb{P}[Z_1 \in 2\mathbb{Z}]$ in Lemma 6.1(ii) yields (3.18). □

Lemma 3.7. *Let $X_1, X_2, \ldots$ be an i.i.d. sequence of $\mathbb{Z}$-valued random variables, $S_n := X_1 + \cdots + X_n$, and $(x_n)_{n\in\mathbb{N}}$ a sequence of even integers. Suppose that $\mathbb{E}[\mathrm{e}^{tX_1}] < \infty$ for some $t > 0$, $\mathbb{P}[X_1 = 0] \wedge \mathbb{P}[X_1 = 2] > 0$, and $x_n = n\mathbb{E}[X_1] + o(n)$. Then*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}[S_n = x_n] = 0.$$

Proof. Write $\mu = \mathbb{E}[X_1]$ and $\sigma^2 = \mathbb{V}\mathrm{ar}[X_1]$, which is finite and positive by the conditions in the lemma. Write $Y_i = X_i - n\mu$ and $y_n = n^{-1/2}\sigma^{-1}(x_n - n\mu)$, so $\mathbb{E}[Y_i] = 0$, $\mathbb{E}[Y_i^2] = \sigma^2$, and $y_n = o(n^{1/2})$. If $Z_n = n^{-1/2}\sigma^{-1}\sum_{i=1}^n Y_i$, Richter's local central limit theorem [18, Chapter 7, §§1 and 4] tells us that

$$\mathbb{P}[S_n = x_n] = \mathbb{P}[Z_n = y_n] = \Theta\left(n^{-1/2} \exp\left\{-\frac{1}{2}y_n^2\left(1 + O(n^{-1/2}y_n)\right)\right\}\right),$$

which is $\exp\{o(n)\}$, since $y_n = o(n^{-1/2})$. □

Second proof of Proposition 2.6. From Lemma 3.6, $n^{-1}\log\pi_n(m)$ can be expressed as

$$n^{-1}\log\mathbb{P}[Z_1^E + \cdots + Z_n^E = m] - n^{-1}\log\mathbb{P}[Z_1 + \cdots + Z_n = m] + \log\cosh\mu - \mu. \tag{3.19}$$

By assumption, $m = m_n$ is such that $m_n/n \to \alpha > 0$. The proof proceeds by choosing μ so that $\mathbb{E}[Z_1^E] = \alpha$; then the first term in (3.19) vanishes in the limit by Lemma 3.7, and the proof of the theorem then reduces to evaluating the other logarithmic rate.

We choose μ so that $\mathbb{E}[Z_1^E] = \alpha$; by Lemma 6.1(ii) this means $\mu\tanh\mu = \alpha$ so that $\mu = \lambda$. Since $Z_1 + \cdots + Z_n$ is Poisson with mean $n\lambda$,

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\log\mathbb{P}[Z_1 + \cdots + Z_n = m_n] &= \lim_{n\to\infty}\frac{1}{n}\log\frac{(n\lambda)^{m_n}\mathrm{e}^{-n\lambda}}{m_n!} \\ &= -\lambda + \lim_{n\to\infty}\frac{m_n}{n}\left[\log n\lambda - \frac{1}{m_n}\log m_n!\right].\end{aligned}$$

Stirling's formula implies that $n^{-1}\log n! = \log(n/\mathrm{e}) + o(1)$, so that

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\log\mathbb{P}[Z_1 + \cdots + Z_n = m_n] &= -\lambda + \lim_{n\to\infty}\frac{m_n}{n}\left[\log\left(\frac{n\lambda\mathrm{e}}{m_n}\right) + o(1)\right] \\ &= -\lambda + \alpha\left(\log\lambda - \log\alpha + 1\right).\end{aligned} \tag{3.20}$$

Combining (3.20) with the $\mu = \lambda$ case of (3.19) we complete the proof. □

4 Proofs of main results

4.1 Exact formula for the expected number of null vectors

Let $\mathcal{N}(n,m;\ell)$ denote the number of left null vectors of weight ℓ, so that

$$\mathcal{N}(n,m) = \sum_{\ell=0}^{m} \mathcal{N}(n,m;\ell). \tag{4.1}$$

The value of $\mathcal{N}(n,m)$ is the number of collections of rows of $M(n,m)$ which sum to $\mathbf{0}$ (mod 2), and for each set of ℓ rows the probability that it sums to $\mathbf{0}$ is $\mathbb{P}_{\rho_n}[A(n,\ell)]$. Hence

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m;\ell)] = \binom{m}{\ell}\mathbb{P}_{\rho_n}[A(n,\ell)], \text{ and} \tag{4.2}$$

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)] = \sum_{\ell=0}^{m}\binom{m}{\ell}\mathbb{P}_{\rho_n}[A(n,\ell)]. \tag{4.3}$$

We can thus express $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$ using our exact formulae for $\mathbb{P}_{\rho_n}[A(n,\ell)]$ given in Lemma 3.1. The proofs of our main results, Theorems 2.7, 2.2, and 2.3, will be based on an asymptotic analysis of (4.3). As in the proof of Theorem 2.5 (see Section 2.3) it is most convenient to work in the binomial model, for which W_n^{bin} is the number of odd components in a multinomial $(W; n^{-1},\dots,n^{-1})$ vector. Thus a key step in the proof will be showing that, in the general case of $W_n \xrightarrow{d} W$, the expression in (4.3) can be well approximated by the binomial case. First, in the next section, we make some preliminary computations.

4.2 Preliminaries

Before embarking on the main proof, we study the rate functions that will appear. Define

$$F_{\rho,\alpha}(\gamma) := \log\left(\frac{(1+\rho(1-2\gamma))^{\alpha}}{2\gamma^{\gamma}(1-\gamma)^{1-\gamma}}\right), \tag{4.4}$$

and recall from (2.1) and (2.2) that $F_\rho(\alpha) = \sup_{\gamma\in[0,1/2]} F_{\rho,\alpha}(\gamma)$ and $\alpha_\rho^* = \inf\{\alpha \geq 0 : F_\rho(\alpha) > 0\}$. Note that for $\gamma \in [0,1/2]$, $\rho(1-2\gamma) \geq 0$. By continuity, $F_{\rho,\alpha}(\gamma)$ attains its supremum over $\gamma \in [0,1/2]$; we denote by $\gamma_0 := \gamma_0(\alpha) \in [0,1/2]$ the *smallest* point at which the supremum is attained.

We collect results on $F_\rho(\alpha)$ and α_ρ^* in the next lemma, which will enable us to complete the proof of Proposition 2.1.

Lemma 4.1. *Suppose that $\mathbb{P}[W = 0] = 0$. For any $\alpha \geq 0$, $F_\rho(\alpha) \geq 0$, and F_ρ is continuous and nondecreasing. The threshold α_ρ^* enjoys the following properties.*

(i) $\alpha_\rho^* \in [0,1]$, and $F_\rho(\alpha) = 0$ for $\alpha \leq \alpha_\rho^*$ but $F_\rho(\alpha) > 0$ for $\alpha > \alpha_\rho^*$.

(ii) If $\alpha < \alpha_\rho^*$, then for any $\varepsilon > 0$, $\sup_{\gamma\in[0,(1/2)-\varepsilon]} F_{\rho,\alpha}(\gamma) < 0$.

(iii) If $\alpha > \alpha_\rho^*$, then $\gamma_0(\alpha) \in [0,1/2)$.

(iv) Suppose that $\tilde{W}$ is another $\mathbb{N}$-valued random variable, with $\tilde{\rho}(s) = \mathbb{E}[s^{\tilde{W}}]$, such that $\tilde{\rho}(s) \leq \rho(s)$ for all $s \in [0,1]$. Then $\alpha_{\tilde{\rho}}^* \geq \alpha_\rho^*$.

(v) $\alpha^*_\rho = 0$ *if and only if* $\mathbb{P}[W = 1] > 0$.

(vi) If $\mathbb{P}[W = 2] = 1$, *then* $\alpha^*_\rho = 1/2$.

(vii) If $\mathbb{E}[W] < \infty$, *then* $\alpha^*_\rho < 1$.

Proof. By Lemma 6.3, $\rho(0) = \mathbb{P}[W = 0] = 0$ and $\rho(1) = 1$; hence $F_{\rho,\alpha}(1/2) = 0$ and $F_{\rho,\alpha}(0) = (\alpha - 1)\log 2$, so that $F_\rho(\alpha) \geq (\alpha-1)^+ \log 2 \geq 0$. Since $F_{\rho,\alpha}(\gamma)$ is nondecreasing as a function of $\alpha \geq 0$, Dini's theorem implies continuity of $F_\rho(\alpha)$ as a function of $\alpha \geq 0$.

For part (i), $F_\rho(\alpha) \geq (\alpha - 1)^+ \log 2$ implies that $F_\rho(\alpha) > 0$ for $\alpha > 1$, so that $\alpha^*_\rho \leq 1$. On the other hand, for $\alpha < 1$, $\gamma_0(\alpha) \in (0, 1/2]$, since by continuity there is some neighbourhood of 0 for which $F_{\rho,\alpha}(\gamma) < 0$.

Since $F_{\rho,\alpha}(\gamma)$ is nondecreasing as a function of α, for $\alpha' \geq \alpha$, $F_\rho(\alpha') \geq F_{\rho,\alpha'}(\gamma_0(\alpha)) \geq F_\rho(\alpha)$, i.e., F_ρ is also nondecreasing. Hence $F_\rho(\alpha) > 0$ for $\alpha > \alpha^*_\rho$. Also, the fact that $F_\rho(\alpha) = 0$ for $\alpha < \alpha^*_\rho$ is immediate from the definition of α^*_ρ and the fact that $F_\rho(\alpha) \geq 0$. Then $F_\rho(\alpha^*_\rho) = 0$ by the continuity of F_ρ established above. Thus we obtain part (i).

For part (ii), suppose that $\alpha < \alpha^*_\rho$. Suppose for some $\gamma_0 \in [0, 1/2)$ that $F_{\rho,\alpha}(\gamma_0) \geq 0$. Since $\rho(1 - 2\gamma) > 0$ for $\gamma < 1/2$, $F_{\rho,\alpha}(\gamma_0)$ is strictly increasing in α, so there exists $\alpha' \in (\alpha, \alpha^*_\rho)$ for which $F_{\rho,\alpha'}(\gamma_0) > 0$, contradicting the definition of α^*_ρ. This gives (ii).

For part (iii), suppose that $\alpha > \alpha^*_\rho$. Then $F_\rho(\alpha) > 0$ by part (i) of the lemma; since $F_{\rho,\alpha}(1/2) = 0$, the supremum is attained in $[0, 1/2)$.

For part (iv), we have that for any $\gamma \in [0, 1/2]$, $F_{\rho,\alpha}(\gamma) \geq F_{\tilde\rho,\alpha}(\gamma)$, since $\tilde\rho(1 - 2\gamma) \geq \rho(1 - 2\gamma)$. So $F_\rho(\alpha) \leq F_{\tilde\rho}(\alpha)$ for all $\alpha \geq 0$, and hence $\alpha^*_{\tilde\rho} \geq \alpha^*_\rho$.

For the remaining parts of the lemma we use more detailed properties of the generating function $\rho(s)$ (see Lemma 6.3). For part (v), differentiating in (4.4) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} F_{\rho,\alpha}(\gamma) = -\frac{2\alpha\rho'(1 - 2\gamma)}{1 + \rho(1 - 2\gamma)} + \log\left(\frac{1-\gamma}{\gamma}\right); \tag{4.5}$$

this is well defined at least for $\gamma \in (0, 1)$. At $\gamma = 1/2$ this equates to $-2\alpha\mathbb{P}[W = 1]$, since by Lemma 6.3 $\rho'(0) = \mathbb{P}[W = 1]$ and $\rho(0) = 0$. So if $\mathbb{P}[W = 1] > 0$, $F_{\rho,\alpha}(\gamma)$ is equal to 0 at $\gamma = 1/2$ and its derivative there is negative for any $\alpha > 0$, so that, for any $\alpha > 0$, $F_{\rho,\alpha}(\gamma) > 0$ for some $\gamma < 1/2$. The 'if' part of part (v) follows.

Conversely, suppose that $\mathbb{P}[W = 1] = 0$. Then the previous argument shows that $F_{\rho,\alpha}(1/2) = F'_{\rho,\alpha}(1/2) = 0$, while a calculation shows that $F''_{\rho,\alpha}(1/2) = 4\alpha\rho''(0) - 4$. Hence by continuity there exists $\delta > 0$ such that for $\alpha < \delta$ and $(1/2) - \delta \leq \gamma \leq 1/2$ we have $F''_{\rho,\alpha}(\gamma) \leq -3$. Hence by Taylor's theorem, $F_{\rho,\alpha}(\gamma) \leq 0$ for $\alpha < \delta$ and $(1/2) - \delta \leq \gamma \leq 1/2$. Also, $F_{\rho,\alpha}(\gamma) \to -\log(2\gamma^\gamma(1 - \gamma)^{1-\gamma})$ as $\alpha \to 0$, which is strictly negative apart from at $\gamma = 1/2$. Thus by Dini's theorem, for all α small enough we have $F_{\rho,\alpha}(\gamma) \leq 0$ for $\gamma \leq (1/2) - \delta$. So all together we have shown that $F_{\rho,\alpha}(\gamma) \leq 0$ for all α sufficiently small. Hence $\alpha^*_\rho > 0$ in this case, giving the 'only if' part of (v).

For part (vi), suppose that $\mathbb{P}[W = 2] = 1$, i.e., $\rho(s) = s^2$. In this case, (4.5) has a zero at $\gamma \in [0, 1/2)$ if $\alpha = s(\gamma)$ where

$$s(\gamma) = \frac{1 + (1 - 2\gamma)^2}{4(1 - 2\gamma)} \log\left(\frac{1-\gamma}{\gamma}\right).$$

We claim that $s(\gamma)$ is decreasing on $[0, 1/2)$, with a unique minimum of $s(1/2) = 1/2$. To verify this, we show $s'(\gamma) < 0$ for $\gamma \in [0, 1/2)$, which, after simplification, amounts to

$$\frac{(1 + (1 - 2\gamma)^2)(1 - 2\gamma)}{8\gamma^2(1 - \gamma)^2} > \log\left(\frac{1-\gamma}{\gamma}\right).$$

Setting $z = 1 - 2\gamma$, it suffices to show that $\frac{z}{(1-z^2)^2} > \frac{1}{2}\log\left(\frac{1+z}{1-z}\right)$ for $z \in (0,1]$, which can be verified by term-by-term comparison of the corresponding power series, namely $z + 2z^3 + 3z^5 + \cdots > z + \frac{z^3}{3} + \frac{z^5}{5} + \cdots$. Hence $s(\gamma) = \alpha$ has no solution for $\alpha < 1/2$, in which case the only stationary value of $F_{\rho,\alpha}$ is at $\gamma = 1/2$, necessarily the maximum. Hence $\alpha^*_\rho \geq 1/2$. On the other hand, if $\alpha > 1/2$ then $F''_{\rho,\alpha}(1/2) = 8\alpha - 4 > 0$, while $F_{\rho,\alpha}(1/2) = F'_{\rho,\alpha}(1/2) = 0$, so by Taylor's theorem and continuity there exists $\gamma < 1/2$ with $F_{\rho,\alpha}(\gamma) > 0$. Hence $\alpha^*_\rho = 1/2$, proving (vi).

Finally we prove part (vii). If $\mathbb{E}[W] < \infty$, Lemma 6.3(ii) implies that, as $\gamma \downarrow 0$, $\rho'(1-2\gamma) = \mathbb{E}[W] + o(1)$. Thus the final term on the right-hand side of (4.5) dominates in the $\gamma \downarrow 0$ limit, and there exists $\delta > 0$ such that $\frac{\mathrm{d}}{\mathrm{d}\gamma}F_{\rho,\alpha}(\gamma) \geq \delta$ for all $\gamma \in [0,\delta]$ and all $\alpha \in [0,1]$. Then by an application of the mean value theorem, $F_{\rho,\alpha}(\delta) \geq (\alpha - 1)\log 2 + \delta^2$ for all $\alpha \in [0,1]$. Thus taking $\alpha < 1$ close enough to 1 we see that $F_{\rho,\alpha}(\delta) > 1$, which implies that $\alpha^*_\rho < 1$. □

Proof of Proposition 2.1. Extract the relevant parts of Lemma 4.1. □

4.3 Approximation by the binomial model

In Section 3.4 we showed (in Lemma 3.5) that $\mathbb{P}_{\rho_n}[A(n,m_n)]$ can be well approximated by $\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]$ on the logarithmic scale, provided that $m_n/n \to \alpha$. The following result is an analogous approximation lemma for $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)]$. One could obtain such a result from Lemma 3.5 applied to (4.3), with some work (including dealing separately with terms with $\ell = o(n)$: cf Section 4.4 below). However, it is more convenient to proceed directly, albeit using similar ideas to the proof of Lemma 3.5; in this case we are helped by the fact that $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$ possesses monotonicity properties absent for $\mathbb{P}_{\rho_n}[A(n,m)]$.

Lemma 4.2. *Suppose that $W_n \xrightarrow{d} W$ and $m_n/n \to \alpha > 0$. Then*

$$\lim_{n\to\infty} n^{-1}|\log \mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)] - \log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)]| = 0.$$

Proof. We use a coupling argument, constructing the general model with row weights distributed as $W_n \xrightarrow{d} W$ on the same probability space as the binomial model with row weights distributed as $W_n^{\mathrm{bin}} \xrightarrow{d} W$ (by Lemma 3.3). For any n, we can use a probability space in which, for each row, the weight in each model converges almost surely to a copy of W. Indeed, let $W(1), W(2), \ldots$ be independent copies of W. Using the Skorokhod representation theorem, we may take $W_{n,1}, W_{n,2}, \ldots$ as independent copies of W_n, being the weights of the rows in the general model, such that $W_{n,i} \to W(i)$ almost surely. Also, take $W^{\mathrm{bin}}_{n,i}$ to be the number of odd components in a multinomial $(W(i); n^{-1}, \ldots, n^{-1})$ distribution, so that $W^{\mathrm{bin}}_{n,1}, W^{\mathrm{bin}}_{n,2}, \ldots$ are independent copies of W^{bin}_n and the weights of the rows in the binomial model.

Let $A_n(i) := \{W_{n,i} \neq W^{\mathrm{bin}}_{n,i}\}$. Then for any $\delta > 0$, we may take n large enough so that $\mathbb{P}[A_n(i)] \leq \delta$, uniformly in i. Let $K(n,m) = \sum_{i=1}^m \mathbf{1}_{A_n(i)}$ denote the number of 'bad' rows. Then $K(n,m)$ is stochastically dominated by a $\mathrm{Bin}(m,\delta)$ variable. In particular, for any fixed $\varepsilon > 0$ and any $C < \infty$, standard binomial tail bounds imply that we may take δ small enough, and hence n sufficiently large, so that

$$\mathbb{P}[K(n,m_n) \geq \varepsilon n] \leq \mathbb{P}[\mathrm{Bin}(2\alpha n, \delta) \geq \varepsilon n] \leq \exp\{-Cn\}. \tag{4.6}$$

We claim that each row added to a matrix can increase the number of null vectors by at most a factor of 2; this follows from (1.1) and (1.2). Hence

$$|\log \mathcal{N}(n, m_n) - \log \mathcal{N}'(n, m_n)| \leq K(n, m_n), \tag{4.7}$$

where $\mathcal{N}(n, m_n)$ is the number of null vectors in the matrix with the $W_{n,i}$ and $\mathcal{N}'(n, m_n)$ is the number of null vectors in the matrix with the $W^{\text{bin}}_{n,i}$. In particular, on $\{K(n, m_n) \leq \varepsilon n\}$, the bound in (4.7) is εn. The statement in the lemma follows from (4.6) and (4.7), since $\varepsilon > 0$ and $C < \infty$ were arbitrary. □

4.4 Null vectors consisting of few rows

In the asymptotics of $\mathbb{E}_{\rho_n}[\mathcal{N}(n, m)]$, it turns out that null vectors of *low weight* play a distinct and important role. Recall (4.1). The main result of this section is the following lemma, which exhibits a polynomial growth rate for null vectors of few rows.

Lemma 4.3. *Suppose that there exist $r_0 \geq 3$ and $r_1 < \infty$ such that $\mathbb{P}[r_0 \leq W_n \leq r_1] = 1$ for all n. Suppose that $m_n/n \to \alpha > 0$. Then there exists $\delta > 0$ such that*

$$\sum_{2 \leq \ell \leq \delta n} \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] = O(n^{2-r_0}). \tag{4.8}$$

Remark. The exponent $2 - r_0$ in (4.8) cannot be improved when $\mathbb{P}[W_n = r_0] > 0$, because $\mathbb{E}[\mathcal{N}(n, m_n; 2)]$ is itself of order n^{2-r_0}. Indeed, there are of order m_n^2 weight-2 candidate vectors, and each is null if each of the two corresponding rows have r_0 non-zeros in matching positions, an event of probability of order n^{-r_0}.

Proof of Lemma 4.3. Let $n, \ell \in \mathbb{N}$. Let $R = R(n, \ell)$ denote the 'column range' of the matrix $M(n, \ell)$, that is, the number of columns of degree at least 1. Let us take $k = k(\ell) \in \mathbb{N}$, to be chosen later. We shall estimate $\mathbb{P}_{\rho_n}[A(n, \ell)]$ by considering separately the events $R \leq k$ and $R > k$.

We interpret $M(n, \ell)$ as arising from a random allocation scheme, where for each row we throw balls at random into n urns (columns). If $R \leq k$ then there is some set of k columns, such that all the balls land in these k columns. For each ball, the probability that it lands in one of the first k columns, given that the other balls cast so far for that row all land in the first k columns, is at most k/n. Hence since for each row at least r_0 balls are cast, and we consider ℓ rows here,

$$\mathbb{P}_{\rho_n}[R \leq k] \leq \binom{n}{k} \left(\frac{k}{n}\right)^{\ell r_0} \leq \frac{n^{k - \ell r_0} k^{\ell r_0}}{k!}. \tag{4.9}$$

If $R > k$ then to have $A(n, \ell)$ occur we need to have each of the columns in the range get hit at least twice (i.e., have degree at least 2). Thus if $R > k$ and $A(n, \ell)$ occurs there is a collection of $k + 1$ columns such that each column in the collection gets hit at least twice. Let $B(i)$ be the event that the column i gets hit at least twice. The probability that a particular entry is 1, given the values of up to k other entries in the same row, is at most $r_1/(n - k)$. Hence the union bound yields for $1 \leq j \leq k + 1$ that

$$\mathbb{P}_{\rho_n}[B(j) \mid \cap_{i=1}^{j-1} B(i)] \leq \binom{\ell}{2} \left(\frac{r_1}{n - k}\right)^2,$$

and hence we have

$$\mathbb{P}_{\rho_n}[\cap_{i=1}^{k+1} B(i)] \leq \left(\binom{\ell}{2} \left(\frac{r_1}{n-k} \right)^2 \right)^{k+1}$$

so that by the union bound, provided $k \leq n/2$ we have

$$\mathbb{P}_{\rho_n}[\{R > k\} \cap A(n,\ell)] \leq \binom{n}{k+1} \left(\binom{\ell}{2} \left(\frac{2r_1}{n} \right)^2 \right)^{k+1} \leq \frac{n^{-(k+1)} \ell^{2(k+1)} c_1^{k+1}}{(k+1)!},$$

where we put $c_1 = 2r_1^2$. Combined with (4.9) this gives

$$\mathbb{P}_{\rho_n}[A(n,\ell)] \leq \frac{n^{k-\ell r_0} k^{\ell r_0}}{k!} + \frac{n^{-(k+1)} \ell^{2(k+1)} c_1^{k+1}}{(k+1)!}.$$

We assume $m_n/n \to \alpha \in (0,\infty)$, so for n large enough so that $m_n \leq (1+\alpha)n$ we have for all ℓ, and for $k \leq n/2$, that

$$\begin{aligned} \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] &= \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(\ell, n)] \\ &\leq \left(\frac{((\alpha+1)n)^\ell}{\ell!} \right) \left(\frac{n^{k-\ell r_0} k^{\ell r_0}}{k!} + \frac{n^{-(k+1)} \ell^{2(k+1)} c_1^{k+1}}{(k+1)!} \right). \end{aligned} \tag{4.10}$$

Taking $k = \ell + r_0 - 2$, we obtain for each fixed ℓ that for some constant $c(\ell)$ we have

$$\begin{aligned} \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] &\leq c(\ell)(n^{\ell(1-r_0)+\ell+r_0-2} + n^{\ell-k-1}) \\ &= c(\ell)(n^{(\ell-1)(2-r_0)} + n^{1-r_0}), \end{aligned} \tag{4.11}$$

which is $O(n^{2-r_0})$ for any fixed $\ell \geq 2$.

Fix an integer $K \geq 2$, to be chosen later, and consider $K \leq \ell \leq \delta n$. Now put $k = \ell(r_0 - 1) - \lceil \ell/2 \rceil$. Assume $\delta \leq 1/(2r_0)$; then for $\ell \leq \delta n$ this choice of k satisfies $k \leq n/2$, so that (4.10) remains valid. Also note that, since $r_0 \geq 3$, $k \geq \frac{3\ell}{2} - 1 \geq \ell$ provided $\ell \geq 2$.

By the bound $\mathrm{e}^\ell \geq \frac{\ell^\ell}{\ell!}$ and similar for k, there are constants c_2, c_3, c_4 such that the first term in the right side of (4.10) (i.e. the product of the first factor with the first term in the second factor) is bounded by a constant times

$$\begin{aligned} \frac{n^{\ell(1-r_0)+k} k^{\ell r_0} c_2^\ell}{\ell^\ell k^k} &\leq \frac{n^{-\lceil \ell/2 \rceil} \ell^{r_0 \ell} c_3^\ell}{\ell^\ell \ell^{(r_0-1)\ell - \lceil \ell/2 \rceil}} \\ &= \left(\frac{c_4 \ell}{n} \right)^{\lceil \ell/2 \rceil}, \end{aligned} \tag{4.12}$$

where for the inequality we used the fact that $\ell \leq k$ to replace k^k by ℓ^ℓ in the denominator.

Similarly, there are constants c_5, c_6, c_7 such that the second term in the right side of (4.10) is bounded by a constant times

$$\begin{aligned} \frac{n^{\ell-k-1} \ell^{2\ell(r_0-1)-2\lceil \ell/2 \rceil+2} c_5^\ell}{\ell^\ell (k+1)^{k+1}} &\leq \frac{n^{\ell(2-r_0)+\lceil \ell/2 \rceil} \ell^{\ell(2r_0-2)-2\lceil \ell/2 \rceil} c_6^\ell \ell^2}{\ell^{\ell r_0 - \lceil \ell/2 \rceil + 1} n} \\ &\leq \left(\frac{c_7 \ell}{n} \right)^{\ell(r_0-2)-\lceil \ell/2 \rceil} (\ell/n). \end{aligned}$$

Combining with (4.12), since $r_0 \geq 3$ so $r_0 - 2 \geq 1$ and $\lceil \ell/2 \rceil \leq (\ell/2) + 1$, we can find a constant c_8 such that for $2 \leq \ell \leq n$ we have

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] \leq c_8 \left(\frac{c_8 \ell}{n} \right)^{\ell/2}.$$

By calculus we have that $\left(\frac{c_8 x}{n}\right)^x$ is decreasing in $x \leq n/(c_8 \mathrm{e})$, so provided $\delta \leq 1/(c_8 \mathrm{e})$ the last bound is maximized, over $K \leq \ell \leq \delta n$, at $\ell = K$, so that

$$\sum_{K \leq \ell \leq \delta n} \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] \leq c_8 \delta n \left(\frac{c_8 K}{n} \right)^{K/2},$$

which is $O(n^{2-r_0})$ provided we choose K so that $K/2 \geq r_0 - 1$. □

4.5 Proof of Theorem 2.2

First we prove Theorem 2.2 for the binomial model, i.e., for $\mathbb{P}^{\mathrm{bin}}_{\rho_n}$ on the right-hand side of (4.3), and then use Lemma 4.2. Specifically, we prove the following result.

Lemma 4.4. *Suppose that $\mathbb{P}[W \geq 1] = 1$. Suppose that $m_n/n \to \alpha \in (0, \infty)$ as $n \to \infty$. Then with $F_\rho(\alpha)$ as defined by (2.2),*

$$\lim_{n \to \infty} n^{-1} \log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n, m_n)] = F_\rho(\alpha). \tag{4.13}$$

Proof. From (3.1),

$$\begin{aligned} \sum_{\ell=0}^{m} \binom{m}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, \ell)] &\leq (m+1)(n+1) 2^{-n} \sup_{0 \leq \ell \leq m} \sup_{0 \leq j \leq n} \binom{m}{\ell} \binom{n}{j} |\rho(1 - (2j/n))|^\ell \\ &\leq (m+1)(n+1) 2^{-n} \sup_{\beta \in [0,1]} \sup_{\gamma \in [0,1]} \binom{m}{\beta m} \binom{n}{\gamma n} |\rho(1 - 2\gamma)|^{\beta m}, \end{aligned} \tag{4.14}$$

setting $\binom{n}{x} = 0$ for $x \notin \{0, 1, \ldots, n\}$. Write

$$S_\alpha(\beta, \gamma) := \left(\frac{|\rho(1-2\gamma)|^\beta}{\beta^\beta (1-\beta)^{1-\beta}} \right)^\alpha \left(\frac{1}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right).$$

Taking $m = m_n = O(n)$ in (4.14) and using the first inequality in (6.3), we obtain

$$n^{-1} \log \sum_{\ell=0}^{m_n} \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n, \ell)] \leq O(n^{-1} \log n) + \log \sup_{\beta \in [0,1]} \sup_{\gamma \in [0,1]} S_{m_n/n}(\beta, \gamma). \tag{4.15}$$

For any $B \geq 0$, routine calculus (with a separate argument for $B = 0$) shows that

$$\sup_{\beta \in [0,1]} \left(\frac{B^\beta}{\beta^\beta (1-\beta)^{1-\beta}} \right) = B + 1,$$

with the supremum attained at $\beta = B/(1+B)$, so that from (4.15) we have

$$n^{-1} \log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n, m_n)] \leq O(n^{-1} \log n) + \sup_{\gamma \in [0,1]} \log \left(\frac{(1 + |\rho(1-2\gamma)|)^{m_n/n}}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right). \tag{4.16}$$

Considering the transformation $\gamma \mapsto 1-\gamma$, we see that

$$\sup_{\gamma\in[1/2,1]}\left(\frac{(1+|\rho(1-2\gamma)|)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right) = \sup_{\gamma\in[0,1/2]}\left(\frac{(1+|\rho(2\gamma-1)|)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right) \le \sup_{\gamma\in[0,1/2]}\left(\frac{(1+\rho(1-2\gamma))^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right),$$

since, for $\gamma \in [0,1/2]$, $|\rho(2\gamma-1)| \le \rho(1-2\gamma)$, by Lemma 6.3(iii). Hence from (4.16) we have, with $F_\rho(\alpha)$ as defined at (2.2), $n^{-1}\log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] \le O(n^{-1}\log n) + F_\rho(m_n/n)$. Since $m_n/n \to \alpha$ and $\alpha \mapsto F_\rho(\alpha)$ is continuous (see Lemma 4.1),

$$\limsup_{n\to\infty} n^{-1}\log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] \le F_\rho(\alpha).$$

For the lower bound, we use the fact that

$$\begin{aligned}\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] &\ge \sum_{\ell=0}^{m_n}\binom{m_n}{\ell}\sum_{j=0}^{\lfloor n/2\rfloor} 2^{-n}\binom{n}{j}(\rho(1-(2j/n)))^\ell \\ &\ge 2^{-n}\sup_{\beta\in[0,1]}\sup_{\gamma\in[0,1/2]}\binom{m_n}{\beta m_n}\binom{n}{\gamma n}|\rho(1-2\gamma)|^{\beta m_n},\end{aligned}$$

using the nonnegativity of the appropriate terms for both inequalities. Using the lower bound in (6.4), similarly to above, we obtain that $\liminf_{n\to\infty} n^{-1}\log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] \ge F_\rho(\alpha)$. Hence combining the upper and lower bounds, we obtain (4.13). □

Now we can give the proof of our main result.

Proof of Theorem 2.2. Lemma 4.4 shows that (2.3) holds for the case where $W_n = W_n^{\mathrm{bin}}$, and Lemma 4.2 shows that the result carries over to the general case. For the final statement of the theorem, suppose that $\alpha < \alpha^*_\rho$ and that $\mathbb{P}[r_0 \le W_n \le r_1] = 1$ for some $r_0 \ge 3$ and $r_1 < \infty$. Lemma 4.3 shows that, for a suitable $\delta > 0$,

$$\sum_{\ell=1}^{\delta n}\binom{m_n}{\ell}\mathbb{P}_{\rho_n}[A(n,\ell)] = O(n^{2-r_0}). \tag{4.17}$$

For $\ell \ge \delta n$, we first restrict to the binomial model. Choose $\varepsilon > 0$ so that $(3\varepsilon)^\delta < 2^{-2\alpha}$. By a similar argument to (4.14), but splitting the supremum over j into two parts,

$$\begin{aligned}\sum_{\ell=\delta n}^{m_n}\binom{m_n}{\ell}\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] &\le (m_n+1)(n+1)2^{-n}\sup_{\delta n\le\ell\le m_n}\sup_{j:|j-(n/2)|\le\varepsilon n}\binom{m_n}{\ell}\binom{n}{j}|\rho(1-(2j/n))|^\ell \\ &\quad + (m_n+1)(n+1)2^{-n}\sup_{0\le\ell\le m_n}\sup_{j:|j-(n/2)|>\varepsilon n}\binom{m_n}{\ell}\binom{n}{j}|\rho(1-(2j/n))|^\ell. \end{aligned}\tag{4.18}$$

Similarly to (4.16), the second term on the right-hand side of (4.18) is bounded above by

$$\exp\left\{o(1) + \sup_{\gamma\in[0,(1/2)-\varepsilon]} F_{\rho,m_n/n}(\gamma)\right\},$$

which decays to 0 exponentially fast, by Lemma 4.1(ii), since $m_n/n \to \alpha \in (0,\alpha^*_\rho)$. On the other hand, for $|j-(n/2)| \le \varepsilon n$, we have from Lemma 6.3(i) that $|\rho(1-(2j/n))| \le 3\varepsilon$,

for $\varepsilon > 0$ small enough, so that, since $\ell \geq \delta n$, $|\rho(1-(2j/n))|^\ell \leq (3\varepsilon)^{\delta n}$, so by the choice of ε the first term in the right hand side of (4.18) tends to 0 exponentially fast. Hence

$$\limsup_{n\to\infty} n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] < 0. \tag{4.19}$$

We next deduce a version of (4.19) with $\mathbb{P}_{\rho_n}$ in place of the special case $\mathbb{P}^{\mathrm{bin}}_{\rho_n}$, using Lemma 3.5 once more. To this end, observe first that the $\mathbb{P}_{\rho_n}$-analogue of the sum in (4.19) consists of $O(n)$ nonnegative terms, so is bounded between the largest term and $O(n)$ times that same term, so that

$$n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] = n^{-1} \log \max_{\delta n \leq \ell \leq m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] + O(n^{-1}\log n).$$

By Lemma 3.5, for any $\varepsilon > 0$, there exist $\varepsilon_{n,\ell}$ with $|\varepsilon_{n,\ell}| \leq \varepsilon$, uniformly for ℓ with $\delta n \leq \ell \leq m_n$ and n sufficiently large, such that

$$\binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] = \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] \exp\{\varepsilon_{n,\ell} n\}.$$

So we obtain

$$\limsup_{n\to\infty} n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] < 0,$$

which, combined with (4.17), yields (2.4). □

5 Cores of sparse random hypergraphs

5.1 Hypergraphs and 2-cores

Given a set $\mathcal{V} = \{v_1, \ldots, v_n\}$, whose elements we call *vertices*, a non-empty subset of $\mathcal{V}$ is called a *hyperedge*. Given a collection $\mathcal{E} := (E_i)$ of m hyperedges, we refer to the pair $(\mathcal{V}, \mathcal{E})$ as a *hypergraph*. This hypergraph may be identified with an $m \times n$ matrix A with entries in $\{0,1\}$ (the *incidence matrix* of the hypergraph), having no zero rows, as follows. The entry $a_{i,j}$ of A takes the value 1 if and only if $v_j \in E_i$, in which case we say row i is *incident* to column j, and that hyperedge E_i is incident to vertex v_j, and refer to (E_i, v_j) as an *incidence* of the hypergraph.

The number of hyperedges incident to a vertex v is the *degree* of v. Fix a hypergraph $(\mathcal{V}, \mathcal{E})$. For $\mathcal{F} \subseteq \mathcal{E}$, the set $V(\mathcal{F}) \subseteq \mathcal{V}$ of vertices which are incident to at least one of the hyperedges in $\mathcal{F}$ is called the *vertex span* of $\mathcal{F}$. We identify the hypergraph $(\mathcal{F}, V(\mathcal{F}))$ by the edge subset $\mathcal{F}$ that induces it, and call $\mathcal{F} \subseteq \mathcal{E}$ a *partial hypergraph*. A partial hypergraph $\mathcal{F} \neq \emptyset$ is a *hypercycle* if every vertex v has even degree with respect to $\mathcal{F}$. For an incidence matrix A, a left null vector is the indicator of a hypercycle in the corresponding hypergraph.

Given a hypergraph $(\mathcal{V}, \mathcal{E})$, the *2-core* is defined via the following algorithm:

1. If there exists no vertex of degree one, stop.
2. Otherwise, select an arbitrary vertex of degree one, and delete the unique incident hyperedge; then return to Step 1.

The algorithm terminates, because the partial hypergraphs are decreasing; the terminal partial hypergraph, which does not depend on the arbitrary choices made in Step 2 (see [11, pp. 127–128]), is called the *2-core* of $\mathcal{E}$, denoted $\mathrm{Core}(\mathcal{E})$. Possibly $\mathrm{Core}(\mathcal{E})$ has no hyperedges. Figure 2 illustrates a hypergraph with 19 vertices and 12 hyperedges, of which 4 hyperedges are in the 2-core.

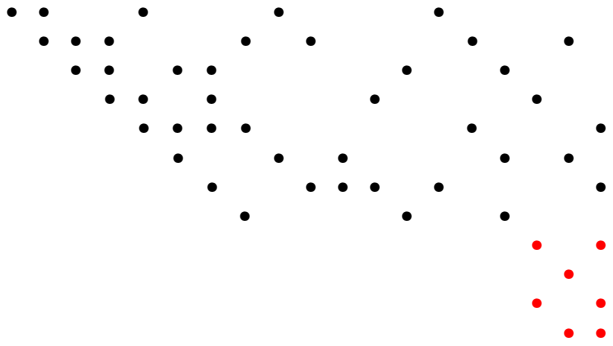

Figure 2: Pictorial representation of the adjacency matrix for a hypergraph with 19 vertices and 12 hyperedges, whose incidences are shown as •. The first 8 hyperedges are not in the 2-core. The last 4 hyperedges form a linear system of rank 3 over GF[2]. The entire 19×12 system has rank 11, and so contains a hypercycle (hyperedges 9 and 11).

The connection between the 2-core and hypercycles was exploited by Cooper [8, p. 371], following an idea that he attributes to Molloy (see [6, p. 268]). The connection is demonstrated by the following useful observation.

Lemma 5.1. *Suppose that the 2-core $\mathcal{C} := \mathrm{Core}(\mathcal{E})$ of a hypergraph $(\mathcal{V}, \mathcal{E})$ has vertex span $V(\mathcal{C}) \subseteq \mathcal{V}$ and size (number of hyperedges) $|\mathcal{C}|$.*

(i) Any hyperedge $E \notin \mathcal{C}$ cannot belong to a hypercycle of $(\mathcal{V}, \mathcal{E})$.

(ii) If $\mathcal{C} = \emptyset$, then $(\mathcal{V}, \mathcal{E})$ contains no hypercycle.

(iii) If $|V(\mathcal{C})| < |\mathcal{C}|$, then $(\mathcal{V}, \mathcal{E})$ contains a hypercycle.

Proof. If there are s hyperedges not in the 2-core $\mathcal{C}$, there exists a labelling of them as $E_1, E_2, \ldots, E_s$ with the property that, for every j, E_j has some vertex with degree one after hyperedges $E_1, E_2, \ldots, E_{j-1}$ are removed. Suppose $(\mathcal{V}, \mathcal{E})$ has some hypercycle $\mathcal{F} \neq \emptyset$. None of $E_1, E_2, \ldots, E_s$ can belong to $\mathcal{F}$: otherwise, there would be some minimum j for which $E_j \in \mathcal{F}$, and this E_j has some vertex v of degree one in the partial hypergraph from which $E_1, E_2, \ldots, E_{j-1}$ have been removed, which contains $\mathcal{F}$; so v cannot have even degree in $\mathcal{F}$, which is a contradiction. This proves (i), and (ii) follows. For (iii), say $c := |V(\mathcal{C})| < |\mathcal{C}| =: r$. Then there are $2^r - 1$ non-empty partial hypergraphs, but only $2^c < 2^r - 1$ possible indicator vectors for a set of vertices of odd degree. By the pigeonhole principle, there must be two distinct partial hypergraphs $\mathcal{F}, \mathcal{F}' \subseteq \mathcal{E}$ for which the sets of vertices of odd degree are the same. Then $\mathcal{F} \triangle \mathcal{F}'$ is a hypercycle. □

5.2 The 2-core in uniform random hypergraphs

In this subsection we consider a certain *uniform random* hypergraph model, which is different from (but related to) the hypergraph model induced by our random matrix

$M(n, m_n)$; in Section 5.3 we will connect the two models. The incidences (non-zero entries) in a random incidence matrix may be viewed as edges of a random bipartite graph, whose left and right nodes are the row labels and column labels, respectively. A standard probability model for a random bipartite graph is to fix the degrees of all nodes in advance (subject to a consistency condition), and sample uniformly from the bipartite graphs with this set of left and right node degrees.

By interpreting an incidence matrix as a hypergraph, as above, this automatically gives a *uniform random hypergraph* model, where the degree of a left node is the *weight* of a hyperedge (meaning its number of incident vertices), and the degree of a right node is the vertex degree defined above. Darling and Norris [12] analyse the statistical properties of the 2-core for such random hypergraphs under suitable conditions on the hyperedge-weight and vertex-degree distributions. In unpublished work of the same authors, a generalization to unbounded vertex degrees and row weights is given, under finite third moments assumptions.

For the purposes of the present paper, we require a more modest relaxation of the conditions of [12], to cover the case where the row weights remain uniformly bounded but the vertex degrees are approximately Poisson distributed.

For each n, define vectors of nonnegative integers $\mathbf{d}_n := (d_n(k) : k \in \mathbb{Z}_+)$ and $\mathbf{w}_n := (w_n(k) : k \in \mathbb{N})$ with $\sum_{k\geq 0} d_n(k) = n$ and $m_n := \sum_{k\geq 1} w_n(k)$; we assume that $\mathbf{d}_n$ and $\mathbf{w}_n$ are compatible in the sense that $\sum_{k\geq 1} k w_n(k) = \sum_{k\geq 0} k d_n(k) < \infty$. We also assume that $m_n \to \infty$. Suppose that for each $i \in \mathbb{N}$ and $j \in \mathbb{Z}_+$,

$$\lim_{n\to\infty} \frac{w_n(i)}{\sum_{k\geq 1} w_n(k)} = \rho_i; \quad \lim_{n\to\infty} \frac{d_n(j)}{n} = \nu_j. \tag{5.1}$$

Define generating functions $\rho(s) := \sum_{k\geq 1} \rho_k s^k$ and $\nu(s) := \sum_{k\geq 0} \nu_k s^k$. We assume that the weights are uniformly bounded, i.e., $\rho_k = 0$ for all w sufficiently large, and the degree distribution has all moments, i.e., $\sum_{k\geq 1} \nu_k k^\beta < \infty$ for all $\beta > 0$. Under these conditions, $\rho'(1)$ and $\nu'(1)$ are the (finite) means corresponding to these distributions.

Consider a sequence of random hypergraphs with n vertices and m_n hyperedges, selected uniformly from those hypergraphs with edge weight multiplicities $\mathbf{w}_n$ and vertex degree multiplicities $\mathbf{d}_n$.

To present asymptotic results for the 2-cores of a sequence of such uniform random hypergraphs, it helps to introduce the notion of sampling a single incidence uniformly at random from all incidences in a hypergraph. Denote such an incidence (E, v). Denote the weight of E by $S+1$, and the degree of v by $L+1$; thus S is the number of other vertices in this hyperedge, and L is the number of other hyperedges incident to this vertex. Size bias occurs here: the event that E has weight k occurs with probability proportional to k times the number of rows of weight k, and similarly the probability that the degree of v is d is proportional to d times the number of degree d vertices. Given the ρ_w and ν_d describing the limiting row weight and vertex degree distributions, we may thus compute a pair of limiting probability generating functions for L and S, respectively:

$$\lambda(s) := \mathbb{E}[s^L] = \sum_{d=0}^{\infty} \lambda_d s^d; \quad \sigma(s) := \mathbb{E}[s^S] = \sum_{w=0}^{\infty} \sigma_w s^w, \tag{5.2}$$

where, due to the size biasing, the coefficients in (5.2) are given by

$$\lambda_d = \frac{(d+1)\nu_{d+1}}{\nu'(1)}; \quad \sigma_w = \frac{(w+1)\rho_{w+1}}{\rho'(1)}.$$

Hence the generating functions themselves become:

$$\lambda(s) = \frac{\nu'(s)}{\nu'(1)}; \quad \sigma(s) = \frac{\rho'(s)}{\rho'(1)}. \tag{5.3}$$

To avoid triviality, we assume that $\sigma_0 = 0$ (equivalently, $\rho_1 = 0$), i.e., there are no 1-edges, and $\lambda_0 \notin \{0, 1\}$ (otherwise the 2-core is of no interest). Define

$$\varphi(s) := 1 - \lambda(1 - \sigma(s)). \tag{5.4}$$

From the conditions $\sigma_0 \neq 1$ and $\lambda_0 \neq 1$, we deduce that $\varphi : [0, 1] \to \mathbb{R}$ is strictly increasing; moreover $\varphi(0) = 1-\lambda(1-\sigma(0)) = 0$ (since $\sigma_0 = 0$) and $\varphi(1) = 1-\lambda_0 \in (0, 1)$, so φ takes values in $[0, 1)$, and there exists a largest solution g^* in $[0, 1)$ of the equation $\varphi(s) = s$. That is,

$$g^* := \sup\{s \in [0, 1) : \varphi(s) = s\}. \tag{5.5}$$

In the case where $g^* > 0$ and the curve $y = \varphi(s)$ crosses the curve $y = s$ (rather than just touching it) at $s = g^*$, we also have

$$g^* = \sup\{s \in (0, 1) : \varphi(s) > s\}. \tag{5.6}$$

Now we can state the result on the 2-core that we shall use, which amounts to a variant of Theorem 7.1 of [12].

Theorem 5.2. *Consider a sequence of uniform random hypergraphs associated with sequences* $\mathbf{w}_n$ *and* $\mathbf{d}_n$ *satisfying (5.1) with* $\rho_w = 0$ *for all* w *large enough and* $\sum_{d\geq 1} \nu_d d^\beta < \infty$ *for all* $\beta > 0$*. Suppose that the corresponding pair (5.2) of random-incidence generating functions has* $\sigma_0 = 0$, $\lambda_0 \notin \{0, 1\}$*, and is such that* g^**, given by (5.5), has either* $g^* = 0$ *or* g^* *satisfying (5.6). Then the following hold a.s. in the limit as* $n \to \infty$*.*

(i) If $g^* = 0$*, the proportion of hyperedges which survive in the 2-core converges to zero.*

(ii) If $g^* > 0$*, then for any* $k \in \mathbb{Z}_+$ *with* $\rho_k > 0$*, the proportion of weight-*k *hyperedges which survive in the 2-core is asymptotically* $(g^*)^k$*; overall, a proportion* $\rho(g^*)$ *of hyperedges survive, and a proportion* $s^*\sigma(g^*)$ *of incidences.*

(iii) If $g^* > 0$*, then for any* $d, k \in \mathbb{N}$ *with* $2 \leq d \leq k$ *and* $\nu_k > 0$*, the proportion of vertices of degree* k *whose degree in the 2-core is* d *converges to*

$$\binom{k}{d} \sigma(g^*)^d (1 - \sigma(g^*))^{k-d}.$$

(iv) If $g^* > 0$*, the 2-core is again a uniform random hypergraph, given its hyperedge weights and vertex degrees, whose distributions are determined by the previous assertions.*

As mentioned above, in [12] all but finitely many coefficients of the generating functions (5.2) were taken to be zero, but the methods admit the modest extension of this section, and indeed can be extended to the case where $\lambda''(1)$ and $\sigma''(1)$ are finite, corresponding to finite third moments for hyperedge weight and vertex degree distributions. Because of its proximity to the result in [12], we do not prove Theorem 5.2 here.

5.3 Application to $M(n, m)$

In the previous subsection we assumed the row and column weights of our random matrix were specified in advance, but now we return to the random matrix model used in the rest of the paper, so our random $m_n \times n$ incidence matrix A will be precisely the matrix $M(n, m_n)$, described in Section 2.1, i.e., with i.i.d. rows with weights having the distribution of W_n, and corresponding generating function $\rho_n(s)$ having limit $\rho(s)$.

To justify being able to apply Theorem 5.2 in this setting, we give the following strong law of large numbers for the empirical distributions of the row and column weights of M.

Lemma 5.3. *Suppose $m_n \in \mathbb{N}$ with $m_n/n \to \alpha$ as $n \to \infty$, with $\alpha > 0$, and the W_n are uniformly bounded. Let $k \in \mathbb{Z}_+$, let $N_k(n)$ be the number of rows of $M(n, m_n)$ of weight k, and let $\tilde{N}_k(n)$ be the number of columns of $M(n, m_n)$ of degree k. Then a.s.,*

$$\lim_{n\to\infty} m_n^{-1} N_k(n) = \mathbb{P}[W = k], \quad \text{and} \tag{5.7}$$

$$\lim_{n\to\infty} n^{-1} \tilde{N}_k(n) = \frac{\mathrm{e}^{-\mu}\mu^k}{k!}, \tag{5.8}$$

where we set $\mu := \alpha\mathbb{E}[W] = \alpha\rho'(1)$. Moreover, the total number of incidences satisfies the law of large numbers

$$\lim_{n\to\infty} n^{-1} \sum_{k\geq 0} k N_k(n) = \lim_{n\to\infty} n^{-1} \sum_{k\geq 0} k \tilde{N}_k(n) = \mu, \text{ a.s.} \tag{5.9}$$

Proof. First note that ${m_n}^{-1}\mathbb{E}[N_k(n)] = \mathbb{P}[W_n = k]$, which converges to $\mathbb{P}[W = k]$ by assumption. To deduce almost sure convergence from this convergence in means, we use the Azuma–Hoeffding inequality in a standard way, as follows. Fix n and for $1 \leq i \leq m_n$ let $\mathcal{F}_i$ be the σ-algebra generated by the rows $X_1, \ldots, X_i$ of $M(n, m_n)$. Define $\xi_i = \mathbb{E}[N_k(n) \mid \mathcal{F}_i]$, with $\xi_0 = \mathbb{E}[N_k(n)]$. Since resampling a single row changes the number of rows of weight k by at most 1, we have for $1 \leq i \leq m$ that

$$|\xi_i - \xi_{i-1}| = |\mathbb{E}[N_k(n) - N_k(n, i) \mid \mathcal{F}_i]| \leq 1,$$

where $N_k(n, i)$ is defined like $N_k(n)$ but based on a matrix with the ith row resampled. By the Azuma–Hoeffding inequality applied to the martingale $(\xi_0, \ldots, \xi_{m_n})$ we obtain for any $\varepsilon > 0$ that

$$\mathbb{P}[|N_k(n) - \mathbb{E}[N_k(n)]| > \varepsilon n] \leq 2\exp(-\varepsilon^2 n/2),$$

so by the first Borel–Cantelli lemma, we have $|N_k(n) - \mathbb{E}[N_k(n)]| \leq \varepsilon n$ for all but finitely many n almost surely. Combined with the convergence of the mean, this gives us (5.7).

The remaining two parts of the lemma use the assumption $\mathbb{P}[W \leq r_1] = 1$ for $r_1 < \infty$. To prove (5.8) note that the weight of the first column (or any other column) of $M(n, m_n)$ is binomially distributed with parameters m_n (number of trials) and $\mathbb{E}[W_n]/n$ (probability of success). Hence $\mathbb{E}[\tilde{N}_k(n)/n] = \mathbb{P}[\mathrm{Bin}(m_n, \mathbb{E}[W_n]/n) = k]$, and by binomial-Poisson convergence this tends to $\mathrm{e}^{-\mu}\mu^k/k!$ as $n \to \infty$. Given this convergence of means, we may prove (5.8) by a similar argument (based on the Azuma–Hoeffding inequality) to the one used to prove (5.7), since resampling a single row changes the number of columns of degree k by at most r_1.

For the final statement in the lemma, we have that

$$n^{-1} \sum_{k\geq 0} k N_k(n) = (m_n/n) \sum_{k=0}^{r_1} k m_n^{-1} N_k(n) \to \alpha \sum_{k=0}^{r_1} k\mathbb{P}[W = k], \text{ a.s.},$$

by (5.7), and then (5.9) follows. □

Corollary 5.4. *Consider the random matrix model $M(n, m_n)$ with row weight distribution $W_n \xrightarrow{d} W$, where the W_n are uniformly bounded. Suppose that $m_n/n \to \alpha > 0$. Then, a.s., taken as hypergraph incidence matrices the sequence $M(n, m_n)$ defines a sequence of uniform random hypergraphs whose row weight and vertex degree distributions satisfy (5.1) with ρ_w and ν_d given by $\rho_w = \mathbb{P}[W = k]$ and $\nu_d = \mathrm{e}^{-\mu}\mu^d/d!$ respectively, where $\mu := \alpha\mathbb{E}[W]$.*

Proof. Since the distribution of $M(n, m)$ is invariant under permutations of the rows or columns, conditional on the empirical distribution of row and column weights, all possible outcomes with those row and column weight distributions are equally likely, so this conditional distribution is indeed uniform. Moreover by Lemma 5.3 the limiting proportion of rows of weight k is given by $\mathbb{P}[W = k]$ and the limiting proportion of columns of degree k is given by $\mathbb{P}[D = k]$ where $D \sim \mathrm{Po}(\mu)$, with $\mu := \alpha\rho'(1)$. Hence conditionally on this sequence of empirical distributions, almost surely we have a sequence of random matrices satisfying the hypotheses of Section 5.2. □

In the notation of Section 5.2, in this case $\nu(s) = \sum_{d=0}^{\infty} \mathrm{e}^{-\mu}\frac{(s\mu)^d}{d!} = \mathrm{e}^{\mu(s-1)}$ is the generating function of a $\mathrm{Po}(\mu)$ random variable, so, by (5.3), the pair (5.2) becomes

$$\lambda(s) = \mathrm{e}^{\mu(s-1)}; \quad \sigma(s) = \frac{\rho'(s)}{\rho'(1)}.$$

In this case, we have from (5.4) that

$$\varphi(s) := 1 - \lambda(1 - \sigma(s)) = 1 - \mathrm{e}^{-\mu\sigma(s)} = 1 - \mathrm{e}^{-\alpha\rho'(s)},$$

and to emphasize the dependence on α we will use the notation φ_α for φ from now on. Recall from (5.5) that g^* was defined as the largest $s \in [0, 1)$ for which $\varphi_\alpha(s) = s$. In order for the model of this section to fit into the setting discussed in Section 5.2, we need to assume that $\sigma_0 = 0$ and $\lambda_0 \notin \{0, 1\}$. Here $\lambda_0 = \mathrm{e}^{-\mu} = \mathrm{e}^{-\alpha\mathbb{E}[W]}$ and $\sigma_0 = \frac{\mathbb{P}[W=1]}{\mathbb{E}[W]}$. So it suffices to assume that $\alpha > 0$, $\mathbb{P}[W \geq 2] = 1$, and $\mathbb{E}[W] < \infty$; in this case the argument in Section 5.2 shows that g^* is well defined.

Note that g^* depends both on ρ and on α; in this section we write $g^* = g^*(\alpha)$ to emphasize the dependence on α; we will show (see Lemma 5.5) that the present definition is equivalent to that at (2.9) given in Section 2.2. For any solution $s \in [0, 1)$ to $\varphi_\alpha(s) = s$, so in particular for $s = g^*(\alpha)$, provided $\rho'(s) \neq 0$, we have $\alpha = h(s)$ as given by (2.7).

We note some facts about $g^*(\alpha)$; recall the definition of $\alpha_\rho^\sharp$ from (2.8).

Lemma 5.5. *Suppose that $\mathbb{P}[W \geq 2] = 1$ and $\mathbb{E}[W] < \infty$. With the convention $\sup\emptyset = 0$, the definition (2.9) is equivalent to the definition (5.5) of $g^*(\alpha)$ as the largest solution of $\varphi_\alpha(s) = s$. Also, $\alpha_\rho^\sharp \in [0, 1]$, and $g^*(\alpha) = 0$ for all $\alpha \in [0, \alpha_\rho^\sharp)$, and for $\alpha > \alpha_\rho^\sharp$, the function $g^*(\alpha)$ is positive and strictly increasing, with $g^*(\alpha) \uparrow 1$ as $\alpha \to \infty$.*

Now assume also that $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W^2] < \infty$. Then the following hold.

(i) We have $g^*(\alpha_\rho^\sharp) \in (0, 1)$ and $\alpha_\rho^\sharp = h(g^*(\alpha_\rho^\sharp)) \in (0, \infty)$.

(ii) The function g^* is right continuous, and there is a finite set $\mathcal{D}_\rho \subset (0, \infty)$, with $\alpha_\rho^\sharp = \inf \mathcal{D}_\rho$, such that g^* is continuous apart from jumps at points of $\mathcal{D}_\rho$. For each $\alpha \in \mathcal{D}_\rho$, $h(g^*(\alpha)) = \alpha$ is a local minimum for h.

(iii) If $\alpha \notin \mathcal{D}_\rho$, then $g^*(\alpha)$ satisfies the crossing condition (5.6).

Proof. Since $\mathbb{P}[W \geq 2] = 1$ and $\mathbb{E}[W] < \infty$, we have $\varphi_\alpha(1) < 1$ and $\varphi_\alpha(0) = 0$. Therefore by continuity we may rewrite (5.5) as

$$g^*(\alpha) = \sup\{s \in (0,1) : \varphi_\alpha(s) \geq s\},$$

using the convention $\sup \emptyset = 0$. By the definition (2.7) of h, for $s \in (0,1)$ it is easy to check that $\varphi_\alpha(s) \geq s$ if and only if $h(s) \leq \alpha$, and this shows that (2.9) and (5.5) give equivalent definitions of $g^*(\alpha)$.

By (2.7) and subsequent remarks, $h(x)$ is positive, continuous in x, and tends to infinity as $x \uparrow 1$. By the definition (2.8), and the subsequent remark, $\alpha_\rho^\sharp \in [0,1]$. By the definition (2.9), it is clear that $g^*(\alpha) = 0$ for $\alpha \in [0, \alpha_\rho^\sharp)$, and the fact that $g^*(\alpha)$ is positive and strictly increasing for $\alpha \in (\alpha_\rho^\sharp, \infty)$ is easily deduced from the continuity of h. Also, given $\varepsilon \in (0,1)$ we can choose α with $h(1-\varepsilon) < \alpha$ so that $g^*(\alpha) > 1 - \varepsilon$, and together with the monotonicity of g^* this shows $g^*(\alpha) \to 1$ as $\alpha \to \infty$.

For part (i), under the extra assumption $\mathbb{P}[W \geq 3] = 1$ we have h going to infinity at 0 and at 1, and by continuity h attains its infimum on $(0,1)$, so using (2.8) and (2.9) we have that $g^*(\alpha_\rho^\sharp)$ is the supremum of a non-empty compact set contained in $(0,1)$, and so lies in (0,1). The last part of (i) also follows from the continuity of h.

For part (ii), under the extra assumption $\mathbb{E}[W^2] < \infty$, note first that if $0 \leq y < \alpha_\rho^\sharp$ then $g^*(y) = 0$. Hence g^* is continuous at y for all $y < \alpha_\rho^\sharp$.

Now let $y \geq \alpha_\rho^\sharp$; note that by (2.9) and continuity of h, we have $h(g^*(y)) = y$. Take a monotonic sequence y_n tending to y; set $x_n = g^*(y_n)$.

Suppose first that $y_n \downarrow y$. Then the sequence x_n is nonincreasing; denoting the limit by x_∞ we have $h(x_n) = y_n$ so $h(x_\infty) = y$ by continuity, and therefore $x_\infty \leq g^*(y)$ by (2.9). Since also $x_n \geq g^*(y)$ by monotonicity we have $x_\infty = g^*(y)$; hence g^* is right-continuous at y.

Now suppose instead that $y_n \uparrow y$. Set $x = g^*(y)$. If h does not have a local minimum at x then $\liminf g^*(y_n) \geq x$, so that $x_n \to x$, and hence g^* is left-continuous at y. Hence, if g^* is discontinuous at y then h has a local minimum at $g^*(y)$.

The function h' is analytic and non-constant on $(0,1)$ so its zeros do not accumulate except possibly at 0 or 1. However $h'(x) = 0$ implies $\rho'(x)/\rho''(x) = -(1-x)\log(1-x)$, so by the assumption $\mathbb{E}[W^2] < \infty$ there exists $\varepsilon > 0$ such that $h'(x) \neq 0$ for $1 - \varepsilon < x < 1$ and for $0 < x < \varepsilon$; for the latter case we use the fact that, as $x \downarrow 0$,

$$\frac{\rho''(x)}{\rho'(x)}(1-x)\log(1-x) \to r_0 - 1 > 1,$$

if $r_0 \geq 3$ is the smallest possible value of W. Thus h has only finitely many local minima in $(0,1)$, and hence h has a local minimum at $g^*(y)$ for at most finitely many y. This completes the proof of (ii).

For part (iii) note that, for $s \in (0,1)$, $\varphi_\alpha(s) > s$ if and only if $h(s) < \alpha$, so (5.6) gives $g^*(\alpha) = \sup\{s \in (0,1) : h(s) < \alpha\}$, which for $\alpha \notin \mathcal{D}_\rho$ agrees with the definition (2.9). □

To apply the results in Section 5.2 to $M(n, m_n)$, we need to assume that $g^*(\alpha)$ either is zero or satisfies (5.6). Lemma 5.5 shows that a sufficient condition for this is that $\alpha \in (0, \infty)$, $\alpha \notin \mathcal{D}_\rho$.

By Theorem 5.2(ii) and (5.9), n^{-1} times the number of incidences which survive in the 2-core converges a.s. to

$$\mu g^* \sigma(g^*) = (\alpha \rho'(1)) g^* \frac{\rho'(g^*)}{\rho'(1)} = \alpha g^* \rho'(g^*) = -g^* \log(1 - g^*). \tag{5.10}$$

By Theorem 5.2(iii) and (5.8), for $d \geq 2$ the proportion of original vertices whose degree in the 2-core is d is asymptotically

$$\sum_{k \geq d} \mathrm{e}^{-\mu} \frac{\mu^k}{k!} \binom{k}{d} \sigma(g^*)^d (1 - \sigma(g^*))^{k-d} = \mathrm{e}^{-\mu} \frac{(\mu\sigma(g^*))^d}{d!} \sum_{j \geq 0} \frac{\mu^j (1 - \sigma(g^*))^j}{j!}$$
$$= \mathrm{e}^{-\mu\sigma(g^*)} \frac{(\mu\sigma(g^*))^d}{d!}, \tag{5.11}$$

the remainder having degree 0 in the 2-core (some columns in the core have degree zero because in the algorithm of Section 5.1, we never delete any columns). In other words, the 2-core vertex degrees have the distribution of a random variable $D\mathbf{1}\{D \neq 1\}$, where $D \sim \mathrm{Po}(\mu\sigma(g^*))$; by (5.10), $\mu\sigma(g^*) = \alpha\rho'(g^*)$. As a check on the previous calculation (5.10) of the number of surviving incidences, n^{-1} times the total number of incidences in the 2-core should converge to the mean of the vertex-degree distribution, which is $\alpha\rho'(g^*)(1 - \mathrm{e}^{-\alpha\rho'(g^*)}) = \alpha g^* \rho'(g^*)$, as in (5.10).

The key issue in predicting hypercycles and rank deficiency is whether the number of rows in the 2-core exceeds the number of occupied columns in the 2-core, which is treated in Theorem 5.6; a related result appears in [8]. Recall the definition of $\psi(g)$ from (2.6) and $\underline{\alpha}_\rho$ from (2.10).

Theorem 5.6. *Suppose W_n are uniformly bounded and $\mathbb{P}[W \geq 3] = 1$. Let $\alpha \in (0, \infty)$. Consider the 2-core of the random incidence matrix $M(n, m_n)$ where $m_n/n \to \alpha$ as $n \to \infty$. Then if $\alpha < \alpha_\rho^\sharp$, the number of rows in the 2-core is $o(n)$, a.s.*

Now suppose $\alpha > \alpha_\rho^\sharp$, so $g^* = g^*(\alpha) > 0$, and suppose that $\alpha \notin \mathcal{D}_\rho$. Then:

(i) n^{-1} times the number of rows in the 2-core converges a.s. to $\alpha\rho(g^*)$.

(ii) n^{-1} times the number of occupied columns in the 2-core converges a.s. to $1 - \mathrm{e}^{-\nu}(1 + \nu)$, where $\nu := \alpha\rho'(g^*)$.

(iii) Almost surely, for all n large enough, the 2-core has more rows than occupied columns if $\psi(g^*(\alpha)) < 0$ but has fewer rows than occupied columns if $\psi(g^*(\alpha)) > 0$. Moreover, there exists $\delta > 0$ such that if $\alpha \in (\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$, for all n large enough, the 2-core has more rows than columns and so the corresponding hypergraph has a hypercycle.

Figure 3 shows an example of some of the more exotic behvaiour that can occur in the random weight setting. In the case where $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$, $\psi(g^*(\alpha))$ changes sign several times, and so Theorem 5.6 shows that as α increases from 0 the 2-core switches from having asymptotically more columns than rows to having more rows than columns not just once (at $\underline{\alpha}_\rho$), but *twice*. This non-monotone behaviour does not occur in the fixed weight case. Proposition 5.8 below, and the subsequent discussion, explains some of the features in the figure.

Proof of Theorem 5.6. By Corollary 5.4, a.s. we have a sequence of random matrices satisfying the hypotheses of Theorem 5.2. If $\alpha < \alpha_\rho^\sharp$, then $g^*(\alpha) = 0$, and Theorem 5.2 shows that the 2-core has $o(n)$ rows. So from now on suppose that $\alpha > \alpha_\rho^\sharp$, so $g^* = g^*(\alpha) > 0$ (see Lemma 5.5).

For the statement (i), note that out of $m_n \sim \alpha n$ rows, a proportion $\rho(g^*)$ survives, by Theorem 5.2(ii). For (ii), the discussion around (5.11) implies that the proportion of

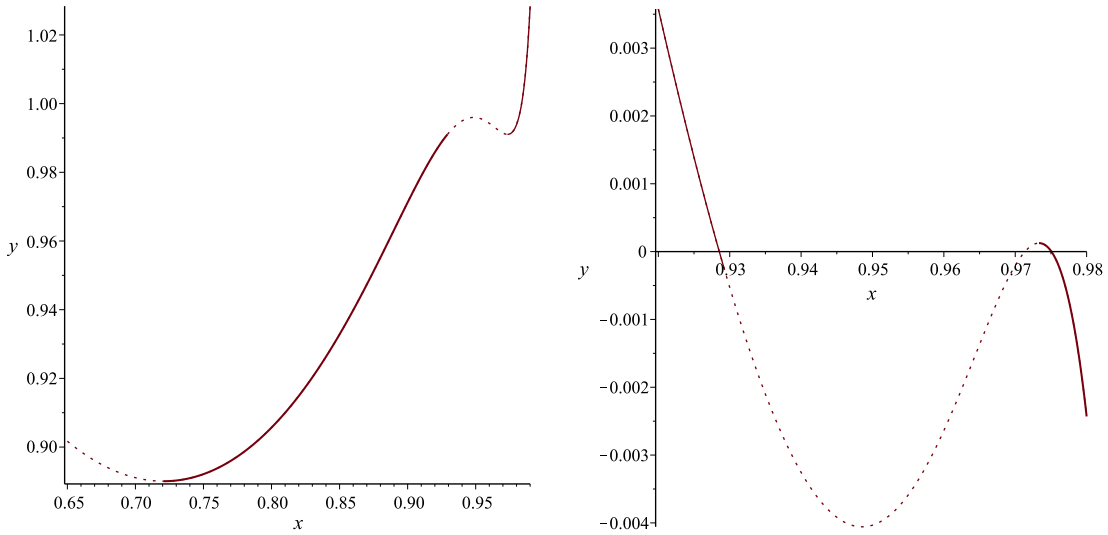

Figure 3: Example with $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$. The left plot shows parts of the curves $y = h(x)$ (all the line) and $x = g^*(y)$ (solid line). The right plot shows parts of the curves $y = \psi(x)$ (dashed line) and the locus of $(g^*(\alpha), \psi(g^*(\alpha)))$ (solid line). Again $g^*(\alpha)$ has two discontinuities, one at $\alpha = \alpha_\rho^\sharp \approx 0.890061$ and one at $\alpha \approx 0.991044$, with the first corresponding to a jump from $g^* = 0$ to $g^* \approx 0.720793$ and the second to a jump from $g^* \approx 0.929269$ to $g^* \approx 0.973325$. The right plot shows the three positive roots of $\psi(x) = 0$. The two of these roots achieved by $\psi(g^*(\alpha))$ are at $x \approx 0.928538$ and $x \approx 0.975069$. The first corresponds to $\alpha = \underline{\alpha}_\rho \approx 0.990686$ and the second to $\alpha \approx 0.991185$. Hence as α ranges in $(0, 1)$, $\psi(g^*(\alpha))$ changes sign from positive, to negative, to positive, and finally to negative again.

the n original vertices whose degree in the 2-core is non-zero is obtained by subtracting from 1 the mass that a $\mathrm{Po}(\nu)$ random variable places on $\{0, 1\}$.

For (iii), we compare the limits in (i) and (ii). Suppose that these limits satisfy

$$\alpha\rho(g^*) > 1 - \mathrm{e}^{-\alpha\rho'(g^*)}(1 + \alpha\rho'(g^*)). \tag{5.12}$$

By our assumptions on α and W, we have $\rho(0) = \rho'(0) = 0$ and $g^* > 0$, which implies that $\rho(g^*)$ and $\rho'(g^*)$ are both positive. Then we may rewrite (5.12) as

$$\begin{aligned} \alpha\rho(g^*) &> 1 - (1 - g^*)(1 + \alpha\rho'(g^*)) \\ &= g^* + (1 - g^*)\log(1 - g^*), \end{aligned}$$

using the definition of g^*. Now substituting in $\alpha = h(g^*(\alpha))$ for α on the left-hand side of the last display (given $\rho'(g^*) > 0$) we may rewrite the last inequality as $\psi(g^*) < 0$, where ψ is defined by (2.6). Similarly, $\psi(g^*) > 0$ is equivalent to

$$\alpha\rho(g^*) < 1 - \mathrm{e}^{-\alpha\rho'(g^*)}(1 + \alpha\rho'(g^*)). \tag{5.13}$$

If $\psi(g^*) < 0$, then (5.12) holds and the limit in (i) in strictly greater than the limit in (ii), which shows that the 2-core eventually has more rows than occupied columns, and vice versa if $\psi(g^*) > 0$ (so that (5.13) holds).

Recall the definition of $\underline{\alpha}_\rho$ from (2.10). We know from Lemma 5.5(iii) that g^* has only finitely many discontinuities. Either $\underline{\alpha}_\rho$ is a continuity point for $g^*(\alpha)$ (and hence for $\psi(g^*(\alpha))$), or else $\underline{\alpha}_\rho \in \mathcal{D}_\rho$ with $\psi(g^*(\underline{\alpha}_\rho)) < 0$ and no other point of $\mathcal{D}_\rho$ is in a

neighbourhood of $\underline{\alpha}_\rho$. In either case, $\psi(g^*(\alpha)) < 0$ for α in an interval of the form $[\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$ with $\delta > 0$. The final statement in the theorem, about the existence of a hypercycle, follows from Lemma 5.1(iii). □

We next state a result giving an upper bound for $\underline{\alpha}_\rho$.

Proposition 5.7. *Suppose that* $\mathbb{P}[W \geq 3] = 1$ *and* $\mathbb{E}[W^2] < \infty$. *Then* $\underline{\alpha}_\rho \leq 1$.

Proof. We know from Lemma 5.5 that $\alpha_\rho^\sharp \leq 1$, so if $\underline{\alpha}_\rho \leq \alpha_\rho^\sharp$ there is nothing to prove. Hence we assume $\underline{\alpha}_\rho > \alpha_\rho^\sharp$ from now on. First we show that

$$\text{for any } \varepsilon > 0 \text{ there exists } \alpha \in (\underline{\alpha}_\rho - \varepsilon, \underline{\alpha}_\rho), \text{ such that } \psi(g^*(\alpha)) > 0. \tag{5.14}$$

By the definition (2.10) of $\underline{\alpha}_\rho$, and the assumption $\underline{\alpha}_\rho > \alpha_\rho^\sharp$, if (5.14) fails then there exists $\delta > 0$ such that $\psi \circ g^*$ is identically zero on the interval $I := (\underline{\alpha}_\rho - \delta, \underline{\alpha}_\rho)$, and by taking δ small enough we may assume the interval I contains no discontinuities of g^*. But then the image $J := g^*(I)$ is also an open interval because g^* is continuous and strictly increasing on I. So we would then have ψ identically zero on J, which would contradict the fact that ψ is analytic and non-constant on $(0, 1)$. Thus (5.14) must hold as asserted.

Observe next that every time the 2-core algorithm deletes a row, it has to create at least one column of degree zero, and possibly more. So the aspect ratio (i.e., number of rows divided by number of occupied columns) is nondecreasing at each step of the algorithm, provided the initial aspect ratio is at least 1. Hence the aspect ratio of a non-empty 2-core is at least as large as the aspect ratio of the original incidence matrix to which the algorithm is applied, provided the latter is at least 1.

So if $m_n/n \to \alpha > 1$, then the aspect ratio of the original matrix exceeds 1 for all n large enough, and hence so does the aspect ratio of the 2-core, assuming it exists. Suppose that $\underline{\alpha}_\rho > 1$. Then by (5.14) and the finiteness of $\mathcal{D}_\rho$, there exists $\alpha' \in (1, \underline{\alpha}_\rho) \backslash \mathcal{D}_\rho$ such that $\psi(g^*(\alpha')) > 0$. Then, by Theorem 5.6(iii), with $m_n/n \to \alpha = \alpha'$, the 2-core has aspect ratio less than 1 for all n large enough, which contradicts the previous conclusion that $\alpha > 1$ implied the 2-core having limiting aspect ratio greater than 1. Hence $\underline{\alpha}_\rho \leq 1$. □

Next we give more information on the key functions h and ψ, which should clarify the situation in Theorem 5.6(iii). By a *root* of ψ, we mean any number x with $\psi(x) = 0$.

Proposition 5.8. *Suppose that* $\mathbb{P}[W \geq 3] = 1$ *and* $\mathbb{E}[W^2] < \infty$. *Then* $0 < \alpha_\rho^\sharp \leq \underline{\alpha}_\rho \leq 1$. *The function* ψ *has at least one root in* $(0, 1)$, *and* h *has at least one local minimum in* $(0, 1)$. *Suppose that the following condition holds:*

(a) h *has a single local minimum* $x_\rho^\sharp$ *in* $(0, 1)$, *with* $h(x_\rho^\sharp) = \inf_{x \in (0,1)} h(x)$.

Then $x_\rho^\sharp$ *is the location of the unique local maximum of* ψ *in* $(0, 1)$, $\psi(x_\rho^\sharp) > 0$, *and the interval* $(0, 1)$ *contains exactly one root of* ψ, *denoted* x_ρ^*, *which satisfies* $x_\rho^\sharp < x_\rho^*$. *Moreover,* $\underline{\alpha}_\rho = h(x_\rho^*) > \alpha_\rho^\sharp$, *and*

$$\psi(g^*(\alpha)) \begin{cases} > 0 & \text{for all } \alpha \in (\alpha_\rho^\sharp, \underline{\alpha}_\rho) \\ < 0 & \text{for all } \alpha > \underline{\alpha}_\rho. \end{cases} \tag{5.15}$$

Finally, in the fixed row-weight case where where $W = r \geq 3$ *a.s., condition (a) holds, and the unique positive root of* ψ *is* $x_r^* \in (\frac{r-2}{r-1}, 1)$.

An important observation that helps to explain the close connection between the functions h and ψ (apparent in Figure 1, for example) and will also form an ingredient in the proof of Proposition 5.8 is the following result.

Lemma 5.9. *For all $x \in (0,1)$, $\psi'(x)$ has the same sign as $-h'(x)$, so, in particular, the locations of the local minima of h correspond exactly to the locations of the local maxima of ψ in $(0,1)$. Finally, if $\mathbb{E}[W^2] < \infty$, then as $x \downarrow 0$ we have*

$$\begin{aligned}\psi(1-x) &= 1 - h(1-x) - x - \frac{\mathbb{E}[W(W-1)]}{2\mathbb{E}[W]}(1+o(1))x^2 \log x \\ &= 1 - h(1-x) - x + o(x). \end{aligned} \tag{5.16}$$

Proof. Differentiating (2.6), we obtain

$$\psi'(x) = -\frac{\rho(x)}{\rho'(x)}\left(\frac{1}{1-x} + \frac{\rho''(x)}{\rho'(x)}\log(1-x)\right). \tag{5.17}$$

On the other hand, from (2.7), we have that, for $x \in (0,1)$,

$$h'(x) = \frac{1}{\rho'(x)}\left(\frac{1}{1-x} + \frac{\rho''(x)}{\rho'(x)}\log(1-x)\right) = -\frac{1}{\rho(x)}\psi'(x),$$

by comparison with (5.17). Finally, (5.16) follows from a routine calculation. □

Before completing the proof of Proposition 5.8, we make some further remarks and present some examples. The main complication in the interpretation of Theorem 5.6(iii) is due to the fact that g^* has discontinuities, so $\{\psi(g^*(\alpha)) : \alpha \geq 0\}$ is only a subset of $\{\psi(x) : x \in [0,1)\}$. Let

$$\mathcal{G}_\rho := \{g^*(\alpha) : \alpha \geq \alpha_\rho^\sharp\}.$$

By Lemma 5.5(ii), $\mathcal{G}_\rho$ is a union of finitely many intervals $\mathcal{G}_\rho = [g_1^-, g_1^+) \cup \cdots \cup [g_\ell^-, g_\ell^+)$ where $g_1^- < g_1^+ < g_2^- < \cdots < g_\ell^+$, and, for each k, $g_k^- = g^*(\alpha)$ for $\alpha \in \mathcal{D}_\rho$, and $h(g_k^-)$ is a local minimum. Recall that $\alpha = h(g^*(\alpha))$ and $\alpha \mapsto g^*(\alpha)$ is increasing for $\alpha > \alpha_\rho^\sharp$ (see Lemma 5.5), so $x \mapsto h(x)$ must be increasing on $\mathcal{G}_\rho$. So in fact $\alpha_\rho^\sharp = h(g_1^-) < \cdots < h(g_\ell^-)$. The 'curve' $\psi(g^*(\alpha))$, $\alpha \geq \alpha_\rho^\sharp$ is then a (discontinuous) trace of $\psi(x)$, where x runs over $\mathcal{G}_\rho$, piecewise continuously on intervals starting at g_k^- which, by Lemma 5.9, correspond to local *maxima* of ψ. Figures 1 and 3 give some illustrations of possible behaviour. Observe that $\psi(x)$ is not necessarily decreasing for all $x \in \mathcal{G}_\rho$.

Note that condition (a) in Proposition 5.8 is not necessary for the sharp transition property (5.15) to hold. Two other relevant conditions are:

(b) ψ has a single root in $(0,1)$;

(c) the global minimum of h on $(0,1)$ is the rightmost local minimum.

If $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, then $h(x) \to \infty$ as $x \to 0$ and as $x \to 1$, so (a) $\Rightarrow$ (c), while in the course of the proof of Proposition 5.8 below, we show that (a) $\Rightarrow$ (b) as well. We mention 3 illustrative examples.

- An example for which conditions (a) and (b) do not hold but (c) does is provided by $\rho(s) = 0.9s^3 + 0.1s^{38}$, for which ψ has 3 positive roots (see Figure 4).

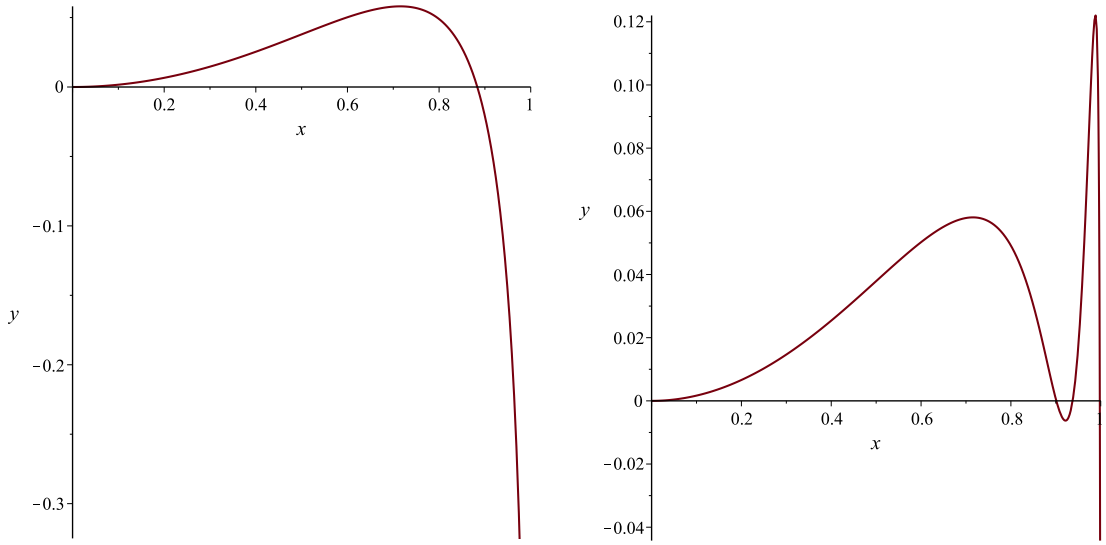

Figure 4: Plots of $y = \psi(x)$ for $\rho(s) = s^3$ (*left*) and $\rho(s) = 0.9s^3 + 0.1s^{38}$ (*right*). In the first case the only positive root is $x_1 \approx 0.883414$, while in the second case the 3 positive roots are $x_1 \approx 0.901174$, $x_2 \approx 0.937414$, and $x_3 \approx 0.997979$. For the case on the right $\alpha_\rho^\sharp \approx 0.872923$ and $g^*(\alpha_\rho^\sharp) \approx 0.988192$, and only the root x_3 exceeds this value. Proposition 5.8 gives $\underline{\alpha}_\rho \approx 0.917935$ for the case on the left and $\underline{\alpha}_\rho \approx 0.998263$ for the case on the right.

- An example for which condition (b) holds but conditions (a) and (c) do not is $\rho(s) = 0.9s^3 + 0.1s^{24}$, for which g^* has two discontinuities (see Figure 1).
- An example in which none of (a), (b) or (c) holds and where (5.15) fails is provided by $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$ (see Figure 3).

Proof of Proposition 5.8. First we show that if $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, then ψ has at least one root in $(0, 1)$. So suppose that there exists an integer $r \geq 3$ for which $\mathbb{P}[W \geq r] = 1$ and $\mathbb{P}[W = r] = p > 0$. Then $\rho(s) \sim ps^r$ as $s \downarrow 0$. From (5.17) we have

$$\psi''(x) = (1-x)^{-1}\left(\frac{2\rho(x)\rho''(x)}{\rho'(x)^2} - 1\right) - (1-x)^{-2}\frac{\rho(x)}{\rho'(x)} - \left(\frac{\rho''(x)}{\rho'(x)} + \frac{\rho(x)\rho'''(x)}{\rho'(x)^2} - \frac{2\rho(x)\rho''(x)^2}{\rho'(x)^3}\right)\log(1-x). \tag{5.18}$$

Taking $x \downarrow 0$ in (5.17) and (5.18), using $\rho^{(k)}(x) \sim \frac{r!}{(r-k)!}px^{r-k}$ for $k \leq 3$, we obtain

$$\psi'(0) = 0; \quad \psi''(0) = \frac{r-2}{r} > 0,$$

since $r \geq 3$. Hence $\psi(0) = 0$ is a local minimum, and $\psi(x) > 0$ for $x > 0$ small enough. But $\psi(x) \to -\infty$ as $x \uparrow 1$, so continuity implies that ψ has at least one root in $(0, 1)$.

Consider the condition (a) in the proposition. Suppose that h has a unique local minimum located at $x_\rho^\sharp \in (0, 1)$, so $\underline{\alpha}_\rho = h(x_\rho^\sharp)$. Then by Lemma 5.9, ψ has a unique local maximum at $x_\rho^\sharp$, and necessarily $\psi(x_\rho^\sharp) > 0$. By continuity (and Rolle's theorem) it follows that ψ has exactly one root $x_\rho^* \in (x_\rho^\sharp, 1)$. So (a) $\Rightarrow$ (b). Moreover, it follows that

$\psi(x) > 0$ for $x \in (0, x^*_\rho)$ and $\psi(x) < 0$ for $x > x^*_\rho$. Hence in this case $\mathcal{G}_\rho = [x^\sharp_\rho, 1)$, and the claim (5.15) follows.

Finally, we show that if $\rho(s) = s^r$ for some $r \geq 3$, then (a) holds. In the case $\rho(s) = s^r$ we obtain (cf (5.18))

$$\psi''(x) = \frac{1}{r(1-x)}\left(r - 1 - \frac{1}{1-x}\right).$$

Hence $\psi''(0) = \frac{r-2}{r} > 0$ and for $x \in (0,1)$ we have $\psi''(x) = 0$ if and only if $x = \frac{r-2}{r-1}$ (an inflexion point). So, by Rolle's theorem, $\psi'(x) = 0$ for at most one $x \in (0,1)$, necessarily $x \in (\frac{r-2}{r-1}, 1)$, and this must be a local maximum for ψ since $\psi(x) \to -\infty$ as $x \uparrow 1$. Another application of Rolle's theorem shows that ψ has a single positive root, x^*_r say, which must be in $(\frac{r-2}{r-1}, 1)$. By Lemma 5.9, the fact that ψ has a single local maximum also shows that h has a single local minimum in $(0,1)$, which is lies in $(\frac{r-2}{r-1}, 1)$. □

Now we can complete the proof of Theorem 2.3.

Proof of Theorem 2.3. The expected number $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$ of null vectors is at least one, and may be large even when $\mathbb{P}_{\rho_n}[T_n \leq \alpha n]$ is small. Nevertheless we can derive bounds on T_n/n by studying the asymptotics of $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$ because

$$\mathbb{P}_{\rho_n}[T_n \leq m] = \mathbb{P}_{\rho_n}[\mathcal{N}(n,m) \geq 2] \leq \mathbb{E}_{\rho_n}[\mathcal{N}(n,m)] - 1, \tag{5.19}$$

by Markov's inequality applied to the nonnegative random variable $\mathcal{N}(n,m) - 1$.

Suppose that $m_n/n \to \alpha \in (0, \alpha^*_\rho)$. Then by (5.19) with (2.4), $\mathbb{P}_{\rho_n}[T_n \leq m_n] = O(n^{-1})$. It follows that, for any $\varepsilon > 0$, $\mathbb{P}_{\rho_n}[T_n \leq (\alpha^*_\rho - \varepsilon)n] \to 0$. On the other hand, Theorem 5.6(iii) implies that there exists $\delta > 0$ such that for any $\alpha \in (\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$, $\mathbb{P}_{\rho_n}[T_n \leq \alpha] \to 1$. Moreover, these results together show that $\alpha^*_\rho \leq \underline{\alpha}_\rho$, and $\underline{\alpha}_\rho \leq 1$ by Proposition 5.7. □

To conclude this section, we give the proof of Proposition 2.8.

Proof of Proposition 2.8. Take $\rho(s) = s^r$ for $r \geq 3$. As already mentioned, the asymptotic for α^*_r is in [5]. Fix $\alpha > 0$. Then,

$$h(1 - \mathrm{e}^{-\alpha r/2}) = -\frac{\log(\mathrm{e}^{-\alpha r/2})}{r(1-\mathrm{e}^{-\alpha r/2})^{r-1}} = \frac{\alpha}{2}(1 + o(1)),$$

as $r \to \infty$. Hence $h(1 - \mathrm{e}^{-\alpha r/2}) \leq \alpha$ for all r sufficiently large, which by (2.8) shows that $\limsup_{r\to\infty} \alpha^\sharp_r \leq \alpha$. Since $\alpha > 0$ was arbitrary, it follows that $\lim_{r\to\infty} \alpha^\sharp_r = 0$.

Finally, by Proposition 5.8, (2.18) holds. Then with (2.22) and repeated Taylor expansions we obtain

$$\begin{aligned}
\underline{\alpha}_r &= -\frac{\log(\mathrm{e}^{-r} + r^2\mathrm{e}^{-2r} + O(r^4\mathrm{e}^{-3r}))}{r(1 - \mathrm{e}^{-r} - r^2\mathrm{e}^{-2r} + O(r^4\mathrm{e}^{-3r}))^{r-1}} \\
&= \frac{1 - r^{-1}\log(1 + r^2\mathrm{e}^{-r} + O(r^4\mathrm{e}^{-2r}))}{(1 - \mathrm{e}^{-r} - r^2\mathrm{e}^{-2r} + O(r^4\mathrm{e}^{-3r}))^{r-1}} \\
&= \left(1 - r\mathrm{e}^{-r} + O(r^3\mathrm{e}^{-2r})\right)\left(1 + (r-1)\mathrm{e}^{-r} + O(r^3\mathrm{e}^{-2r})\right) \\
&= 1 - \mathrm{e}^{-r} + O(r^3\mathrm{e}^{-2r}),
\end{aligned}$$

completing the proof of (2.21). □

6 Technical appendix

6.1 Parity of random variables

The following simple result, on the probability that certain integer-valued random variables take even values, will be used several times. The formula in Lemma 6.1(i) for the probability that a binomial variable is even may be found for example in [16, pp. 277–278].

Lemma 6.1. *Let X be a $\mathbb{Z}_+$-valued random variable with probability generating function $\phi(s) := \mathbb{E}[s^X]$. Then $\mathbb{P}[X \in 2\mathbb{Z}] = \frac{1}{2}(1 + \phi(-1))$. In particular (i) if $X \sim \mathrm{Bin}(n, p)$, then $\mathbb{P}[X \in 2\mathbb{Z}] = (1 + (1-2p)^n)/2$; and (ii) if $X \sim \mathrm{Po}(\mu)$, then $\mathbb{P}[X \in 2\mathbb{Z}] = \mathrm{e}^{-\mu} \cosh \mu$.*

Proof. For $X \in \mathbb{Z}_+$, $\mathbf{1}\{X \in 2\mathbb{Z}\} = \frac{1}{2}(1 + (-1)^X)$, yielding the first statement in the lemma. For the $\mathrm{Bin}(n, p)$ case, $\phi(s) = (ps + 1 - p)^n$, giving part (i), while in the $\mathrm{Po}(\mu)$ case, $\phi(s) = \mathrm{e}^{\mu(s-1)}$, which gives $\mathbb{P}[X \in 2\mathbb{Z}] = (1 + \mathrm{e}^{-2\mu})/2 = \mathrm{e}^{-\mu} \cosh \mu$. □

6.2 Parity of multinomial random variables

We saw in Lemma 6.1 that the probability that a $\mathrm{Bin}(n, p)$ random variable is even is $(1+(1-2p)^n)/2$. In this section we extend this formula to a more complicated multinomial setting and more general congruence conditions modulo r.

Here is our probabilistic model. We perform a sequence of n independent trials. Each trial is probabilistically identical, and we are interested in the outcome of a trial described in terms of an arbitrary collection of k events $A_1, \ldots, A_k$. Probabilities $p(E_I)$ are specified for each of the 'elementary' events E_I defined as

$$E_I := (\cap_{i \in I} A_i) \cap \left(\cap_{j \in [k] \setminus I} A_j\right).$$

for each $I \subseteq [k]$ (here $[k] := \{1, \ldots, k\}$). We assume that $p(E_I) > 0$ for each I and that $\sum_{I \subseteq [k]} p(E_I) = 1$. Use $\mathbb{P}_n$ to denote the probability measure associated with the model consisting of n trials as just described.

Let N_i be the number of occurrences of event A_i. For $L \subseteq J \subseteq [k]$ set

$$E_{J,L} := (\cap_{i \in L} A_i) \cap \left(\cap_{i \in J \setminus L} A_i^c\right)$$

and set $E_J := E_{[k],J}$. Also, set $p_o(J)$ to be the probability for a single trial that an odd number of outcomes $A_i, i \in J$ occur, and note that

$$1 - 2p_o(J) = \sum_{L \subseteq J} (-1)^{|L|} p(E_{J,L}),$$

where $|L|$ denotes the number of elements of L. The next result gives a general formula for the probability in n trials that for each i the N_i falls into a particular congruence class modulo r, and a specialization to the event that N_i is even for each i in a specified subset I of $[k]$ and N_i is odd for each $i \in [k] \setminus I$. For positive integer r, define the complex number $\omega := \mathrm{e}^{(2\pi/r)i}$ (a complex rth root of unity).

Lemma 6.2. *(i) Let $r \geq 2$ be an integer and $\mathbf{t} = (t_1, \ldots, t_k) \in \{0, 1, \ldots, r-1\}^k$. Then*

$$\mathbb{P}_n\left[\cap_{i=1}^k \{N_i \equiv t_i \pmod r\}\right] = r^{-k} \sum_{\mathbf{h} \in \{0,1,\ldots,r-1\}^k} \omega^{-\mathbf{t}\cdot\mathbf{h}} \left(\sum_{\mathbf{g} \in \{0,1\}^k} \omega^{\mathbf{g}\cdot\mathbf{h}} p_{\mathbf{g}}\right)^n, \qquad (6.1)$$

where for $\mathbf{g} = (g_1, \ldots, g_k) \in \{0,1\}^k$ *we set* $p_{\mathbf{g}} = p(E_{\mathbf{g}})$ *with the event*

$$E_{\mathbf{g}} := \left(\cap_{i\in\{1,\ldots,k\}:g_i=1} A_i\right) \cap \left(\cap_{i\in\{1,\ldots,k\}:g_i=0} A_i^c\right)$$

defined in terms of a single trial, and $\mathbf{t}\cdot\mathbf{h} := \sum_{i=1}^k t_i h_i$ *and* $\mathbf{g}\cdot\mathbf{h} := \sum_{i=1}^k g_i h_i$.

(ii) For any $I \subseteq [k]$,

$$\mathbb{P}\left[(\cap_{i\in I}\{N_i \in 2\mathbb{Z}\}) \cap \left(\cap_{j\in[k]\setminus I}\{N_j \notin 2\mathbb{Z}\}\right)\right] = 2^{-k}\sum_{J\subseteq[k]} (-1)^{|J\cap I|}(1-2p_o(J))^n. \tag{6.2}$$

Proof. Part (ii) follows from the $r = 2$ case of part (i) on setting t_i to be 1 on I and 0 otherwise. Thus we need to prove part (i). We obtain (6.1) by induction on n. First consider the case $n = 0$. In this case the left side of (6.1) is equal to 1 if $t_i = 0$ for all i and is equal to zero otherwise, and the right side is equal to

$$r^{-k}\sum_{\mathbf{h}\in\{0,1,\ldots,r-1\}^k} \omega^{-\mathbf{t}\cdot\mathbf{h}}.$$

If $\mathbf{t} = \mathbf{0}$ then this expression is 1. Otherwise, it is zero since if $t_i \neq 0$ for some i then

$$\sum_{h_i=0}^{r-1} \omega^{-t_i h_i} = \frac{1-\omega^{-rt_i}}{1-\omega^{-t_i}} = 0.$$

Thus the inductive hypothesis holds for $n = 0$. Suppose it holds for some n. In the case of $n+1$ trials, conditioning on the outcome of the first trial we obtain

$$\begin{aligned}
\mathbb{P}_{n+1}\left[\cap_{i=1}^k\{N_i \equiv t_i \pmod r\}\right] &= \sum_{\mathbf{g}\in\{0,1\}^k} p_{\mathbf{g}}\mathbb{P}_n\left[\cap_{i=1}^k\{N_i \equiv t_i \pmod r\}|E_{\mathbf{g}}\right] \\
&= \sum_{\mathbf{g}\in\{0,1\}^k} p_{\mathbf{g}}\mathbb{P}_{n+1}\left[\cap_{i=1}^k\{N_i \equiv t_i - g_i \pmod r\}\right] \\
&= r^{-k}\sum_{\mathbf{g}\in\{0,1\}^k} p_{\mathbf{g}} \sum_{\mathbf{h}\in\{0,1,\ldots,r-1\}^k} \omega^{-(\mathbf{t}-\mathbf{g})\cdot\mathbf{h}}\left(\sum_{\mathbf{f}\in\{0,1\}^k}\omega^{\mathbf{f}\cdot\mathbf{h}}p_{\mathbf{f}}\right)^n \\
&= r^{-k}\sum_{\mathbf{h}\in\{0,1,\ldots,r-1\}^k} \omega^{-\mathbf{t}\cdot\mathbf{h}}\left(\sum_{\mathbf{g}\in\{0,1\}^k}\omega^{\mathbf{g}\cdot\mathbf{h}}p_{\mathbf{g}}\right)\left(\sum_{\mathbf{f}\in\{0,1\}^k}\omega^{\mathbf{f}\cdot\mathbf{h}}p_{\mathbf{f}}\right)^n \\
&= r^{-k}\sum_{\mathbf{h}\in\{0,1,\ldots,r-1\}^k} \omega^{-\mathbf{t}\cdot\mathbf{h}}\left(\sum_{\mathbf{g}\in\{0,1\}^k}\omega^{\mathbf{g}\cdot\mathbf{h}}p_{\mathbf{g}}\right)^{n+1},
\end{aligned}$$

which completes the induction. □

6.3 Generating function properties

The next result collects some elementary properties of probability generation functions.

Lemma 6.3. *Let* $\phi(s) := \mathbb{E}[s^X]$, $s \in [-1,1]$, *for a* $\mathbb{Z}_+$*-valued random variable* X. *Then* $\phi(0) = \mathbb{P}[X = 0]$, $\phi(1) = 1$, *and* $\phi(s)$ *is infinitely differentiable at least for* $s \in (-1,1)$*; if* $\mathbb{E}[X] < \infty$ *then* $\phi'(s) = \frac{\mathrm{d}}{\mathrm{d}s}\phi(s)$ *is continuous in the closed interval* $[-1,1]$. *Moreover,*

(i) Suppose that $\mathbb{P}[X=0]=0$. *Then as* $s \downarrow 0$,

$$\phi(s) = s\mathbb{P}[X=1] + O(s^2), \text{ and } \phi'(s) = \mathbb{P}[X=1] + O(s).$$

(ii) If $\mathbb{E}[X] < \infty$, *then as* $s \downarrow 0$,

$$\phi(1-s) = 1 - s\mathbb{E}[X] + o(s), \text{ and } \phi'(1-s) = \mathbb{E}[X] + o(1).$$

(iii) For any $s \in [0,1]$, $|\phi(-s)| \le \phi(s)$.

Proof. Apart perhaps from part (iii), all of the properties stated in the lemma are well known: see for example [16, pp. 264–266]. For part (iii), let $s \in [0,1]$. Then

$$|\phi(-s)| \le \mathbb{E}[\left|(-s)^X\right|] = \mathbb{E}[s^X] = \phi(s).$$

□

6.4 Asymptotic estimates

We shall use the following bounds on the binomial coefficient $\binom{n}{k}$.

Lemma 6.4. *Let* $n \in \mathbb{N}$ *and* $k \in \{0,1,\dots,n\}$. *Then*

$$\binom{n}{k} \le \left(\left(\frac{k}{n}\right)^{k/n} \left(1-\frac{k}{n}\right)^{1-(k/n)} \right)^{-n} \le n^k \mathrm{e}^k k^{-k}. \tag{6.3}$$

On the other hand, if $0 < k < n$,

$$\binom{n}{k} \ge \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \mathrm{e}^{-1/6} \left(\left(\frac{k}{n}\right)^{k/n} \left(\frac{n-k}{n}\right)^{(n-k)/n} \right)^{-n}. \tag{6.4}$$

Proof. We apply Robbins's refinement of Stirling's formula (see e.g. [16, §II.9]), which says that for any $n \ge 1$,

$$n! = (2\pi)^{1/2} n^{n+(1/2)} \mathrm{e}^{-n+\varepsilon_n},$$

where $\frac{1}{12n+1} < \varepsilon_n < \frac{1}{12n}$. This yields the upper bound, for $n \ge 1$ and $k, n-k \ge 1$,

$$\binom{n}{k} \le \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \left(\left(\frac{k}{n}\right)^{k/n} \left(\frac{n-k}{n}\right)^{(n-k)/n} \right)^{-n}, \tag{6.5}$$

where we have used the fact that

$$\varepsilon_n - \varepsilon_k - \varepsilon_{n-k} \le \frac{1}{12n} - \frac{12n+2}{144k(n-k)+12n+1} \le \frac{1}{12n} - \frac{12n+2}{36n^2+12n+1} \le 0,$$

since $k(n-k) \le n^2/4$. By considering separately the cases (i) $k \in \{0,n\}$, and (ii) $0 < k < n$, using (6.5) in case (ii), we obtain the first inequality in (6.3). The second inequality in (6.3) follows from the fact that

$$\left(1-\frac{k}{n}\right)^{-(n-k)} = \left(1+\frac{k}{n-k}\right)^{n-k} \le \mathrm{e}^k.$$

For the lower bound, another application of Robbins's bounds yields (6.4), where for the $\mathrm{e}^{-1/6}$ term we have used the fact that $\varepsilon_n - \varepsilon_k - \varepsilon_{n-k} \ge -\frac{1}{12k} - \frac{1}{12(n-k)} \ge -\frac{1}{6}$. □

Acknowledgements

The authors thank Julian West for pointing out the elementary application of the pigeonhole principle in the proof of Lemma 5.1, thus avoiding the use of linear algebra.

References

[1] R.C. Alamino and D. Saad, Typical kernel size and number of sparse random matrices over Galois fields: A statistical physics approach, *Phys. Rev. E* **77** (2008) 061123.

[2] G.V. Balakin, The distribution of the rank of random matrices over a finite field, *Theory Probab. Appl.* **13** (1968) 631–641.

[3] G.V. Balakin, V.F. Kolchin and V.I. Khokhlov, Hypercycles in a random hypergraph, *Discrete Math. Appl.* **2** (1992) 563–570.

[4] N.J. Calkin, Dependent sets of constant weight vectors in GF(q), *Random Struct. Alg.* **9** (1996) 49–53.

[5] N.J. Calkin, Dependent sets of constant weight binary vectors, *Combinat. Probab. Comput.* **6** (1997) 263–271.

[6] C. Cooper, Asymptotics for dependent sums of random vectors, *Random Struct. Alg.* **14** (1999) 267–292.

[7] C. Cooper, On the distribution of rank of a random matrix over a finite field, *Random Struct. Alg.* **17** (2000) 197–212.

[8] C. Cooper, The cores of random hypergraphs with a given degree sequence, *Random Struct. Alg.* **25** (2004) 353–375.

[9] K.P. Costello and V. Vu, On the rank of random sparse matrices, *Combinat. Probab. Comput.* **19** (2010) 321–342.

[10] R.W.R. Darling, D.A. Levin, and J.R. Norris, Continuous and discontinuous phase transitions in hypergraph processes, *Random Struct. Alg.* **24** (2004) 397–419.

[11] R.W.R. Darling and J.R. Norris, Structure of large random hypergraphs, *Ann. Appl. Probab.* **15** (2005) 125–152.

[12] R.W.R. Darling and J.R. Norris, Differential equation approximations for Markov chains, *Probab. Surv.* **5** (2008) 37–79.

[13] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink, Tight thresholds for cuckoo hashing via XORSAT, *Proc. 37th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, 2010, Volume 6198/2010, Springer, pp. 213–225.

[14] O. Dubois and J. Mandler, The 3-XORSAT threshold, *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002, pp. 769–778 (version dated 28 February 2003).

[15] P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960) 17–61.

[16] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. I, 3rd ed., John Wiley, New York, 1968.

[17] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, 2nd ed., John Wiley, New York, 1971.

[18] I.A. Ibragimov and Yu.V. Linnik, Independent and Stationary Sequences of Random Variables, Walters-Noordhoff, Groningen, 1971.

[19] M. Ibrahimi, Y. Kanoria, M. Kraning, and A. Montanari, The set of solutions of random XORSAT formulae, preprint (2011) `arXiv:1107.5377`.

[20] S. Janson, Poisson convergence and Poisson processes with applications to random graphs, *Stochastic Process. Appl.* **26** (1987) 1–30.

[21] N.L. Johnson and S. Kotz, Urn Models and Their Application: An Approach to Modern Discrete Probability Theory, John Wiley & Sons, New York, 1977.

[22] V.F. Kolchin, Cycles in random graphs and hypergraphs (abstract), *Adv. in Appl. Probab.* **24** (1992) 768.

[23] V.F. Kolchin, Random graphs and systems of linear equations in finite fields, *Random Struct. Alg.* **5** (1994) 135–146.

[24] V.F. Kolchin, Random Graphs, Cambridge University Press, 1999.

[25] V.F. Kolchin, A threshold effect for systems of random equations in finite fields, *Discrete Math. Appl.* **9** (1999) 355–364.

[26] V.F. Kolchin, B.A. Sevastyanov, and V.P. Chistyakov, Random Allocations, V.H. Winston & Sons, Washington, 1978.

[27] I.N. Kovalenko, On the limit distribution of the number of solutions of a random system of linear equations in the class of Boolean functions, *Theory Probab. Appl.* **7** (1967) 47–56.

[28] I.N. Kovalenko and A.A. Levitskaya, Stochastic properties of systems of random linear equations over finite algebraic structures, pp. 64–70 in: Probabilistic Methods in Discrete Mathematics, V.F. Kolchin *et al.* (eds.), TVP/VSP, 1993.

[29] A.A. Levitskaya, Systems of random equations over finite algebraic structures, *Cybernet. Systems Anal.* **41** (2005) 67–93.

[30] H.M. Mahmoud, Pólya Urn Models, CRC Press, 2009.

[31] M. Mézard, G. Parisi, and R. Zecchina, Analytic and algorithmic solution of random satisfiability problems, *Science* **297** (2002) 812–815.

[32] P.A.P. Moran, An Introduction to Probability Theory, OUP, 1968.

[33] T. Muetze, Generalized switch-setting problems, *Discrete Math.* **307** (2007) 2755–2770.

[34] M. Talagrand, Spin Glasses: A Challenge for Mathematicians, Springer-Verlag, Berlin, 2003.